2017

# Subjective beliefs and decision making under uncertainty in the field

Sandip Kumar Agarwal
*Iowa State University*

**Subjective beliefs and decision making under uncertainty in the field**

by

**Sandip Kumar Agarwal**

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:
Quinn Weninger, Major Professor
Bruce Babcock
Otávio Camargo-Bartalotti
Keri L. Jacobs
Rajesh Singh

The student author and the program of study committee are solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2017

## DEDICATION

I would like to dedicate this thesis in the lotus feet of my spiritual mentors, His Holiness Nirankari Baba Hardev Singh Ji and his wife, Her Holiness Satguru Mata Savinder Hardev Ji for their untiring efforts towards making me a better human being. They had been the constant source of inspiration and the ultimate source of divine blessings for me. They have imbibed within me the quotient of Universal Brotherhood through the Fatherhood of God, which has unfettered me from the shackles of ignorance that has created distances among people on the planet. Since my childhood, I have received the blessings from my spiritual mentors in one or the other form, and an opportunity to do a PhD. was one such blessing. I planned to present my thesis to Baba ji upon completion. However, in a play of unfortunate events last year, He left his physical body and merged into Almighty Nirankar (God). Through this thesis I make a humble effort to pay my tributes to His Holiness Babaji, whose life was a tribute to Humanity as he has been love and humanity personified. I take this opportunity to dedicate this thesis to Her Holiness Satguru mataji, thank her and seek her guidance and blessings for the future.

I also dedicate this thesis to my younger brother, Piyush Agarawal (16), who bravely fought cancer and merged into Almighty Nirankar during my PhD. His memories and courage will always be close to my heart and keep inspiring me.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# 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 thesis. First and foremost, Dr. Quinn Weninger for his guidance, patience and support throughout my journey of doctoral research. His insights and words of encouragement have 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: Dr. Keri L. Jacobs, Dr. Otávio Bartalotti, Dr. Rajesh Singh, and Dr. Bruce Babcock. The comments and insights of my committee members have always helped me in improving this thesis. I owe special thanks to Dr. Jacobs for her involvement in my PhD. from its initial stages to in its present form. Her inputs to the present research through numerous discussions have been invaluable. I sincerely thank my major professor, my committee members, Dr. John S. Schroeter and Dr. Brent Kreider for making adequate arrangements for me during an unanticipated family emergency. I sincerely appreciate their co-operative gesture towards my situation. I would also acknowledge the financial support from Center for Agricultural and Rural Development (CARD) and the Department of Economics during my PhD.

My immediate family including my parents and sister, and my extended family including my uncle, aunt and cousins have been extremely supportive throughout my PhD. Their sacrifice while I have been away from them always motivated me towards putting my best. I extend heartfelt thanks to all my friends in and around Ames, who have been like my family during my stay in Ames, and made my stay in Ames memorable. Last but not the least, I would like to thank my wife, Shikha Agarwal, who has traveled with me in this journey. Although, physically we have been in two different countries, she made every possible effort to understand me and develop trust in me. Her every possible effort towards my PhD. and our relationship deserves a thousand thanks !

# ABSTRACT

Nutrient management decision under uncertainty is a critical and complex decision that a farmer has to make on his field. It is complex as it is a decision that may be linked to several other decisions on their field.

Research studies have shown that nutrient application alters the crop yield density. This indicates that nutrient is not limited to be a productive input, but it can also be used as a tool for risk management under uncertainty in agriculture.

Researchers have developed models of decision making under uncertainty in agriculture and elsewhere, where strong restrictions on the decision making agents' perception and preferences has been imposed to identify the underlying decision process. Similar, framework has been used for studying the farmer's nutrient decision (mostly based on expected utility or expected profit maximization framework). In the process of modeling the nutrient decision, farmer's perception (expectations) about the nitrogen uncertainty is artificially constructed, which is assumed by researchers to be *rational expectations*. It is important to note that the choice of optimal nitrogen in a field is a subjective concept, which rests upon the truth and validity of the assumptions introduced in the decision making framework.

This dissertation relaxes those arbitrary assumptions about the nitrogen uncertainty by measuring the subjective uncertainty perceived by a farmer surrounding the chosen level of nitrogen. Although, the uncertainty around the chosen level of nitrogen is measured, nothing much can be said about the choice of optimal nitrogen. The subjective expectations of farmer around the optimal level of nitrogen are measured and juxtaposed with the agronomic benchmark. This dissertation is a contribution to the field by providing factual evidence about the discordance between the subjective beliefs of farmers to objective reality. More broadly, this research is an effort towards advancement of the study of agriculture decision making under uncertainty by measurement of subjective expectations of farmer in context to nitrogen yield

mapping, which when combined with risk preferences of farmer may be able to identify the true underlying nitrogen decision model for a farmer.

# CHAPTER 1. INTRODUCTION

Nutrient management decision under uncertainty is a critical and complex decision that a farmer has to make on his field. It is complex as it is a decision that may be linked to several other decisions on their field.

Research studies have shown that nutrient application alters the crop yield density. This indicates that nutrient is not limited to be a productive input, but it can also be used as a tool for risk management under uncertainty in agriculture.

Researchers have developed models of decision making under uncertainty in agriculture and elsewhere, where strong restrictions on the decision making agents' perception and preferences has been imposed to identify the underlying decision process. Similar, framework has been used for studying the farmer's nutrient decision (mostly based on expected utility or expected profit maximization framework). In the process of modeling the nutrient decision, farmer's perception (expectations) about the nitrogen uncertainty is artificially constructed, which is assumed by researchers to be *rational expectations*. It is important to note that the choice of optimal nitrogen in a field is a subjective concept, which rests upon the truth and validity of the assumptions introduced in the decision making framework.

This dissertation relaxes those arbitrary assumptions about the nitrogen uncertainty by measuring the subjective uncertainty perceived by a farmer surrounding the chosen level of nitrogen. Although, the uncertainty around the chosen level of nitrogen is measured, nothing much can be said about the choice of optimal nitrogen. The subjective expectations of farmer around the optimal level of nitrogen are measured and juxtaposed with the agronomic benchmark. This dissertation is a contribution to the field by providing factual evidence about the discordance between the subjective beliefs of farmers to objective reality. More broadly, this research is an effort towards advancement of the study of agriculture decision making under uncertainty by measurement of subjective expectations of farmer in context to nitrogen yield mapping, which when combined with risk preferences of farmer may be able to identify the true underlying nitrogen decision model for a farmer.

## 1.1 Overview

The dissertation is organized as follows. Chapter 2 outlines the development of the survey instrument that has been used to measure farmers' subjective beliefs about nutrient management on their fields. The chapter develops a background and rationale for the measurement of subjective beliefs. Based on recent literature and methods in psychology and behavioral economics, the survey method is discussed. An outline of the data collection methodology, and summary statistics of the data collected through the survey is provided in this chapter. The chapter also provides inputs and comments on improving the survey method for future waves of this study based on the lesson learned.

The purpose of chapter 3 is to provide an objective benchmark for comparison of the elicited subjective beliefs of farmers in chapter 2. The chapter reviews studies that have modeled ex-post realized yield data to provide information about yield distributions. Based on previous studies and existing econometric development, a generalized linear model with a beta distribution is used to model the moments of the yield distribution conditional on nitrogen and weather variables. The effect of nitrogen and weather on the first three moments of the yield density and the marginal productivity of nitrogen (MPN) is discussed. Some of the research question that this chapter focuses on include

1. Do the yield nitrogen relationship support a plateau response function ?
2. What are the implications of weather uncertainty for the choice of optimal nitrogen ?
3. Is the MPN convex in nitrogen under weather uncertainty ?
4. Do higher nitrogen application increase yield variance under weather uncertainty ?
5. Do higher nitrogen application increase yield skewness under weather uncertainty ?

Chapter 4 uses data and findings from chapters 2 and 3 to build the main results of the dissertation. A multilevel model is used to model the subjective yields of farmers, using the data collected through the survey methodology described in chapter 2. Multilevel models are mostly used in psychology and education literature, and are not common in the economics literature. The modeling technique is discussed in detail with an attempt to develop it parallel to regression analysis, which is familiar to reader in economics. Primarily, the subjective MPN estimates are compared with the objective MPN estimates developed in chapter 2. Some of the research question that this chapter focuses on include

1. How does the subjective MPN estimates of farmer compare with the objective estimates of MPN ?

2. How are the field and farmer characteristics associated with the subjective beliefs of expected yield corresponding to optimal nitrogen application?

3. How does the measure of subjective yield skewness from farmers subjective belief about yield distribution compare with the objective measure of yield skewness ?

Chapter 5 concludes by providing an overview of the contribution of this dissertation. It also discusses the limitations of the present research by clearly outlining what this research can and cannot achieve. This chapter summarizes the results of each chapter and discusses their possible implications. It also outline potential research questions for future research in this direction.

# CHAPTER 2. ELICITATION OF SUBJECTIVE BELIEFS: A PILOT STUDY OF FARMER'S NITROGEN DECISION-MAKING IN CENTRAL IOWA [1]

## 2.1 Introduction

In the summer of 2014 a survey of Iowa crop producers was conducted to (1) learn about common nutrient management practices, and (2) to elicit farmers' subjective beliefs about the weather and crop growth uncertainty they face when making nutrient management decisions. The survey is part of a larger study that seeks to uncover the subjective or *perceived* relationship between nitrogen application practices, e.g., quantity, timing, application method, and crop yield outcomes. This chapter provides a rationale for measuring subjective beliefs in the context of decision making under uncertainty and discusses specific design elements of the 2014 survey instrument. This chapter also provide summary statistics for the information that is gathered and discuss lessons learned from the pilot study.

### 2.1.1 Background

Measurement of subjective beliefs and expectations has been incorporated in studies of school choice and returns to schooling (Dominitz and Manski, 1996; Jensen, 2010; Zafar, 2013; Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2014), subjective income expectations (Dominitz and Manski, 1996, 1997b; Dominitz, 2001), perceptions of economic insecurity (Dominitz and Manski, 1997a; Manski and Straub, 2000; Campbell et al., 2007), subjective health expectations (Delavande, 2008; Delavande and Kohler, 2009, 2012), consumption and investment decisions (Dominitz and Manski, 2004; Gouret and Hollard, 2011), and energy

[1] This chapter is reproduced from the working paper, Agarwal, Sandip; Jacobs, Keri L.; and Weninger, Quinn, "Elicitation of Subjective Beliefs: A Pilot study of farmers' nitrogen management decision-making in Central Iowa" (2016). *Economics Working Papers. 6, Iowa State University*

choices (Blass et al., 2010; Allcott, 2013). Recent developments in elicitation methods have also been successful in eliciting subjective expectations in developing countries where populations are generally less educated and with illiterate respondents (Attanasio, 2009; Delavande et al., 2011a,b; Delavande, 2014).

Belief elicitation methodology is broadly categorized as indirect or direct. Indirect methods use responses to a question or choices based on a task designed to infer a respondents' degree of uncertainty. The most popular indirect methods include the Gambling Method, Bid Method, Lottery Method, Odds Method, Weighting Method, Ranking Method, Visual Counter Method and Smoothing Method. Although indirect methods are believed to be the easiest for respondents to understand and use (Winkler (1967); Chesley (1978)), a common critique is that the interplay of an individual's risk attitude with the perceived uncertainty may influence the response and choices made by the respondent in the assessment of subjective probability (Chesley (1978))[2].

Direct methods measure subjective beliefs by, as the name suggests, asking respondents directly to report the likelihood of outcomes or to assess and report probabilities. Two common methods for measuring subjective probabilities are used. In the first approach probability intervals are elicited, in which respondents are presented with a range of possible outcomes and are asked to report the likelihood that the outcome falls within specified intervals, e.g. the respondent may be asked to report the likelihood that the value of the random variable will lie below a threshold value. This has been referred to as *Percent Chances* format by Manski (2004) and more recently termed as the *Subjective Probability Interval Estimate* or *SPIES* method by Haran et al. (2010), which is the terminology that will be used for the method used in this chapter hereafter. The second approach is the fractile method, which asks respondents to identify one or more points on the support of a subjective distribution that match particular likelihood outcomes, e.g., the respondent may be asked to name the $95^{th}$ percentile value of the support of their subjective likelihood of a random event. Direct methods of probability

[2] For a detailed discussion of these methods and their comparison, please look at (Winkler, 1967; Hampton et al., 1973; Ludke et al., 1977; Chesley, 1978; Norris et al., 1990)

elicitation have the advantage that they do not require the researcher to pre-commit to a particular distribution support or functional forms in analyzing responses.

While some researchers believe the SPIES method performs better than fractile approaches, (Winkler, 1967; Ludke et al., 1977; Chesley, 1978), there is no consensus. Further, evidence is lacking as to which method is cognitively simpler for respondents to use: the fractile method requires only equally likely responses (Chesley, 1978) but may be more difficult than the SPIES method because one-half of the distribution is disregarded (Huber, 1974). Also, it may be difficult for respondents to assess small probabilities using the fractile method because it produces relatively tight subjective distributions, particularly in the tails. The fractile method is subject to error accumulation if the median is elicited inaccurately because all successive fractiles build on the median (Winkler, 1967; Hampton et al., 1973; Schaefer and Borcherding, 1973; Seaver et al., 1978; Alpert and Raiffa, 1982).

An increasingly important issue in the elicitation of subjective beliefs is over-precision, which is a form of overconfidence whereby respondents believe they have more control over random outcomes, or underestimate the range of possible outcomes than is objectively warranted. Recent research has linked the finding of overprecision to the elicitation method (Haran et al., 2010). Under the SPIES approach, respondents are asked to report likelihoods over a wide distribution support. The SPIES method may help the respondent think about all possible outcomes and thereby reduce the inclination to overlook and underweight low-probability outcomes, i.e., outcomes that fall in the tails of the subjective distribution under investigation.

Subjective beliefs of decision-makers under uncertainty plays an important role in agricultural economics research. Attempts to elicit and incorporate farmers' subjective beliefs into models and analysis are fairly rare. Exceptions include studies that measure beliefs about yield loss due to crop disease or adverse weather (Carlson, 1970; Pingali and Carlson, 1985; Menapace et al., 2013), subjective yield, price and income expectations (Grisley and Kellogg, 1983; Clop-Gallart and Juárez-Rubio, 2007), subjective beliefs about optimal nitrogen applications (SriRamaratnam et al., 1987), and beliefs about weather impacts (Sherrick et al., 2000; Sherrick, 2002). These studies have relied on indirect methods of probability elicitation or fractile approaches, an exception being (Sherrick et al., 2000; Sherrick, 2002) which uses both a fractile and inverse CDF approach.

In 2014, a pilot survey instrument was developed and administered to measure the subjective beliefs of farmers regarding the nitrogen management practices they use on their cropped fields. The survey incorporated the latest methodological advances from the above cited literature. The SPIES methodology is employed. Seventy seven farmers participated in the study. Section 2.2 describes the procedures that were used to develop the survey instrument and administration of the survey to Iowa farmers. Summary statistics of survey responses are reported in section 2.3. The lessons that are learned in the pilot study are discussed in section 2.4. Concluding remarks are presented in section 2.5.

## 2.2 Survey Instrument Development and Administration

A draft survey was developed in spring of 2014. A focus group meeting was held with six farmers in February of 2014. The final web-based version of the survey was completed in the summer of 2014. The survey questions are presented in section 2.6 in the Appendix.

Focus group meetings were conducted with local producers and agronomists to ensure survey questions were interpreted as intended and that the coverage of response options was reasonable. The survey was revised based on the feedback. A concern raised in the focus group is the importance of maintaining producers anonymity, particularly because the survey asked producers to reveal production practices. Anonymity of the responses was achieved with the help of the cooperative firm that facilitated the list of potential participants. Participant names and contact information was maintained by the coop firm. The researcher team was provided with number IDs that could not be traced to individual producers. Survey participants received a signed confidentiality agreement by ground mail (the agreement also appeared with the online survey instrument and is shown in an appendix). As is common in survey studies, respondents who completed the survey were compensated for their time (Dominitz and Manski, 1996; Delavande, 2008; Arcidiacono et al., 2012; Zafar, 2013). The coop firm managed the allocation of compensation crediting each producer who completed the survey \$50 on their coop account.

The survey involved human participants. Exemption from requirements of the human subject protections regulations was obtained via the Institutional Review Board (IRB) at the Iowa State University, IRB ID 14-245.

Invitations to participate in an online Qualtrics survey were mailed from the cooperative to approximately 500 of their farmer-members. Each producer used a unique code to start the survey. 96 responses were collected, which is a response rate of 19%.

### 2.2.1 Survey Question Design

The survey includes four main sections. The first and fourth sections gather general information about farm characteristics and respondent demographics. For example, respondents provided information about the scale and scope of the farm operation, the nature of the management process, e.g., whether the respondent makes nutrient management decisions unilaterally or with a management team, the experience level and education of the respondent, and what if any external advice influences management choices. The middle two sections focus on the specific nutrient management practices used on specific fields. The middle sections also measure subjective beliefs about crop production and weather uncertainty.

Discussions with farmers during focus group interviews revealed potentially important differences in nutrient management strategies across fields of varying quality. Therefore, survey was designed so as to assure coverage of varying land qualities. Survey sections two and three repeat the same set of field-specific questions but for different fields. Respondents were first asked, in section 2, to answer questions as they apply to their *Best Producing field.* In the third survey section respondents, via a randomization process, were asked question as they pertain to either an *Average Producing Field* or a *Typically Under-Performing Field.* Conditioning responses on specific fields of varying quality served two purposes. It is believed that asking the farmer to think about a particular field reduced ambiguity, e.g., the survey asked about specific management actions on a particular field rather than general management approaches. Second, the randomization between average and under-performing fields provides variation crucial for identification of differences in management across field quality types, and to identify and quantify heterogeneity in respondents' subjective beliefs about uncertainty as it pertains to fields of varying quality and productivity. Inducing variation in field quality through the randomization mechanism is an important innovation of the survey instrument.

The survey was administered online using Qualtrics software (an overview is available here). Among the Qualtrics' useful features is the ability to condition questions on a specific response to an earlier question and use branching features to customize question based on their relevance. For example, farmers who indicated their farm produced livestock were shown questions specific to their livestock production operations; these questions did not appear to respondents who indicated their farms specialized in crop production. Elicitation of the subjective distribution of random outcomes provides an example of the advantages of question customization features of the Qualtrics technology. Following a series of lead-up questions about management practices respondents were asked to report expected production, in bushels per acre. The value the respondent reported was saved and recalled for questions later in the survey.

Among the variables recorded in sections two and three are field characteristics and production practices, including the county in which the field is located, total acres, Corn Suitability Rating (CSR), proportion of the field that is Highly Erodible Land (HEL), the crop rotation, tillage practices, and nutrient testing. CSR is an index used to measure a field's potential for corn productivity.[3] Highly Erodible Land (HEL) indexes susceptibility to soil erosion and can play an important role in management. Rotation patterns also matter to nutrient decisions, particularly nitrogen for corn production. Corn is a nitrogen-intensive crop whereas soybeans are a legume that fixes nitrogen in soil. Soybeans in rotation with corn are used to manage soil fertility and reduce the need for added nitrogen. Respondents were asked to report their crop in the current year and previous two years (i.e. 2014 through 2012) thus allowing to identify the specific rotation used. Tillage practices indicate the intensity of the production system and nutrient testing may reduce uncertainty about the soil nutrient profile, thus impacting beliefs (Babcock (1992)).

---

[3]For a detailed discussion of CSR, refer to *Corn Suitability Ratings An Index to Soil Productivity*, PM1168, August 2009, Iowa State University Extension Publication

Farm management encompasses a number of interrelated and potentially complicated decisions. Nutrient management decisions, as with most farm management decisions, are made under uncertainty over prices, weather, and other factors. The goal of this study is to understand how farmers perceive uncertainty and how the various sources and extent of uncertainty influence management actions. An important objective of the survey is the measure farmers' subjective beliefs about the nutrient management problem, particularly the decisions regarding nitrogen application on managed fields.[4] The questioning strategy underlying belief elicitation are discussed next.

Farmers were asked to rate the fertility of their field using a 5 point adequacy scale, believing this reflects the respondents' own subjective assessment of how productivity potential that drives management choices. Focus group meetings indicated that it is difficult to precisely estimate the nitrogen concentration in soil. Soil testing provides precise measurement of nitrogen concentration. However, not all farmers conduct soil tests and test results can vary within a field and across time. As a follow-up to the fertility adequacy questions, respondents reported their best estimate of how much nitrogen they felt was required to achieve their expected yield target. Respondents were then asked to indicate how confident they were in their beliefs about nitrogen requirements. Confidence was assessed, for example, by asking respondents to select whether the true soil nitrogen concentration was within 5%, 10%, 20% and 50% of their estimate.

The nutrients available for plant uptake vary substantially with soil type, weather conditions and the plant growth. This introduces non-uniformity in the availability of nutrients, which also depends on the timing of nitrogen application. Respondents were asked to select specific months in which commercial nitrogen and manure application occurred or were planned. Corresponding to each selected month of nitrogen application, respondents were asked to report the quantity of nitrogen in pounds per acre that was applied or is planned for application, the N-P-K ratio, and the method of application.

[4]Emphasis on nitrogen application decisions is further motivated by its importance in reducing nitrogen runoff and improving water quality in agriculture.

### 2.2.2 Measuring expectations: Yield and rainfall

One goal of the survey is to elicit subjective beliefs about randomness in crop production. A series of questions were posed to measure the location and shape of a subjective yield distribution. This question was approached by asking the respondent to report the *Expected Yield* on the managed field under consideration. The response to this question was then used to frame additional questions about expected yield randomness. The approach is to construct four threshold values from the subjective yield distribution support. Two thresholds on either side of the Expected Yield response were generated: 75% and 90% and 110% and 125% of the reported expected yield value. Hereafter, these thresholds are referred to as T1, T2, T3 and T4. It should be noted that T1-T4 are customized to the *Expected Yield* response and generated automatically by the Qualtrics software. By scaling the thresholds in this way, each respondent is asked about yield thresholds on customized subjective distribution support centered on their expected yield.

The respondent was asked to report probabilities that their realized yield will fall within particular intervals. Probabilities were elicited using the *Chances out of 100* format. Figure 2.1 provides a screen shot of the actual questions.

The sequence in which this and similar "chances" questions were presented and the use of a sliding scale for reporting chances are important features of the survey design. The first question posed is "What are the chances out of 100 that the yield will be below T2 bushels per acre." The respondent used a computer mouse to slide a pointer across a 0-100 scale to enter their answer. Any value between 0 and 100 (inclusive) was permitted. The second question posed is "What are the chances out of 100 that the yield will be below T1 bushels per acre." The response was entered using the computer mouse and the same sliding scale.

The Qualtrics software was used to check consistency of a producer's responses, in particular to assure reported values are consistent with the axioms of probability. If a respondent's answer violated the axiom, a warning message displayed. For example, the chances out of 100 that the realized yield is less than T1 cannot exceed the chances out of 100 that it is less than T2. If an entered response violated any of the axioms, a message appeared on the computer

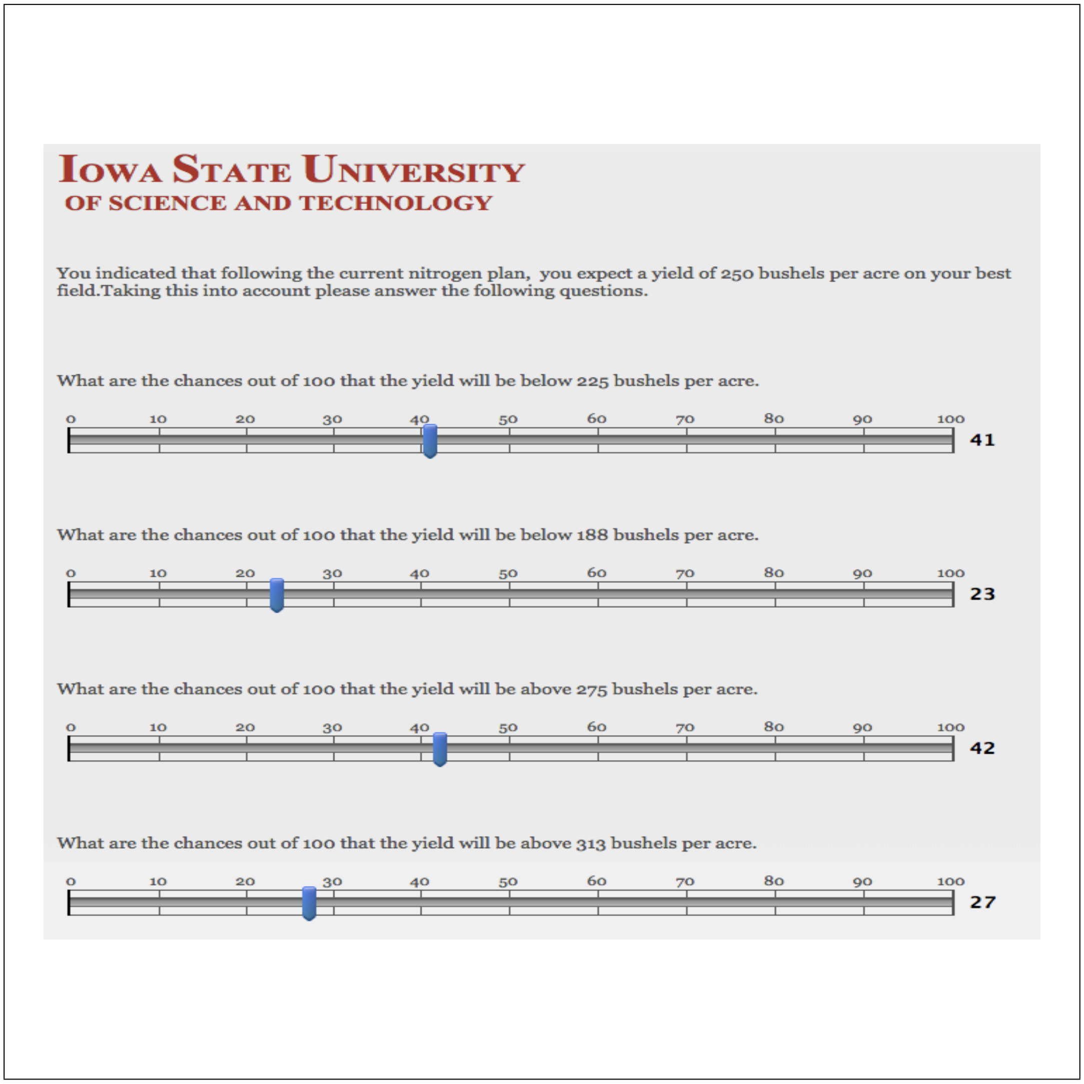



Figure 2.1 Survey Page for Eliciting Subjective Probabilities

screen explaining the error and asked the respondent to review their responses and enter a new chances response. If the new response still violated the axiom of probability, the respondent was allowed to proceed to the next question. We permitted the respondent to move on with additional questions even with a response that was inconsistent believing that a single error message may help guide consistency while multiple error messages could discourage respondents from completing the survey.

Focus group discussion indicated that rainfall in July is an important determinant of crop growth. Farmers' subjective beliefs about rainfall distributions were elicited using a similar *chances out of 100* format, though unlike in the expected yield probability elicitation described above, threshold rainfall levels were anchored on the mean observed rainfall levels in central Iowa. Respondents were told that the mean July rainfall in central Iowa is 4.3 inches. These *chances out of 100* questions used thresholds of 0.5, 2.0, 6.5, 8.5 inches per month.

The survey focus then turned to the relationship between nitrogen and yield. Respondents were reminded about the nitrogen applications plans they revealed earlier in the survey (programming in Qualtrics permitted us to recall prior answers from the respondents). The respondent was asked to report expected yield (in bu./acre) if 115%, 130%, 85% and 75% of their most recent nitrogen application was instead applied. They were also asked to report expected yields conditional on the their planned nitrogen application but under alternative July rainfall scenarios: rainfall levels of 6.5 inches, 8.5 inches, 2 inches and 0.5 inches for the month of July.

## 2.3 Descriptive Statistics

Of the 96 survey responses, 72 were fully completed, 24 contained some missing information, of which 19 were less than 50% complete. These 19 incomplete surveys were deemed unreliable and dropped from the analysis that follows. The data contain the 77 fully or mostly complete responses to questions: 72 complete and 5 incomplete on respondents' *Best producing fields*, 34 responses for *Average producing fields*, and 38 responses for *Under performing fields*. The next subsection summarize the responses and highlight the important findings and opportunities for additional work on this issue.

### 2.3.1 Land ownership and farmer demographics

Table 2.1 provides a summary of the respondents' demographics. *Land farmed* is defined as the aggregate of farmland that is owned, leased, or under alternative tenancy arrangement. Respondents' answers ranged from 43 to 11,000 acres, indicating considerable variation in the size of respondents' farming operations. Among the 77 respondents who answered questions on farm size, 13% report they do not own any land and 39% report they do not farm on leased

Table 2.1 Demographics and Land Ownership.

| | N | Mean | Median | Std. Dev. | Min. | Max. |
|---|---|---|---|---|---|---|
| Land farmed (acres) | 77 | 693.8 | 412 | 1,308 | 43 | 11,000 |
| Owned farmland (%) | 77 | 56.26 | 69 | 41.42 | 0 | 100 |
| Leased farmland (%) | 77 | 40.18 | 26 | 41.1 | 0 | 100 |
| Time spent farming (%) | 71 | 56.61 | 55 | 28.88 | 3 | 100 |
| Age (years) | 71 | 58.23 | 60 | 14.85 | 25 | 91 |
| Experience (years) | 71 | 33.77 | 36 | 15.7 | 1 | 70 |

land, i.e., they own all of the land they farm. The range of respondents' age and experience farming is also noted.

Respondents were asked to report their educational attainment. Of the 71 responses to this question, 41 (57.75%) indicated a high school education, 24 (33.80%) indicated they had earned an undergraduate degree and 6 (8.45%) reported completing graduate education (a M.S. or Ph.D. degree). Among farmers who hold an undergraduate degree or higher, 43.3% report that they share nutrient management decision making responsibilities with one or more business partners. 31.7% of respondents with a high school education share decision making responsibility with others.

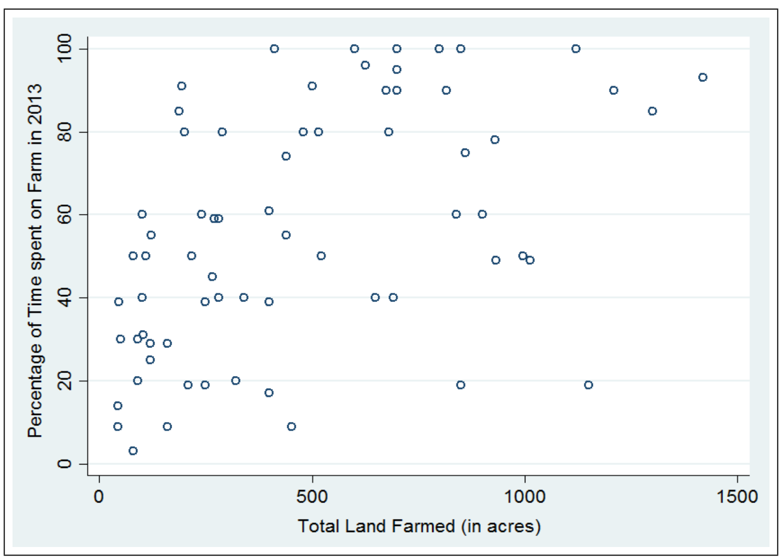



Figure 2.2 Time spent farming versus size of operation.

Midwest U.S. farmers often work at multiple jobs. Figure 2.2 plots the labor hours allocated to farming as a percentage of total labor hours worked (on the vertical axes) against the size of the farm operation (on the horizontal axes). The simple correlation between the labor dedicated to farming and farm size is 0.491 (p-value $< 0.000$).[5]

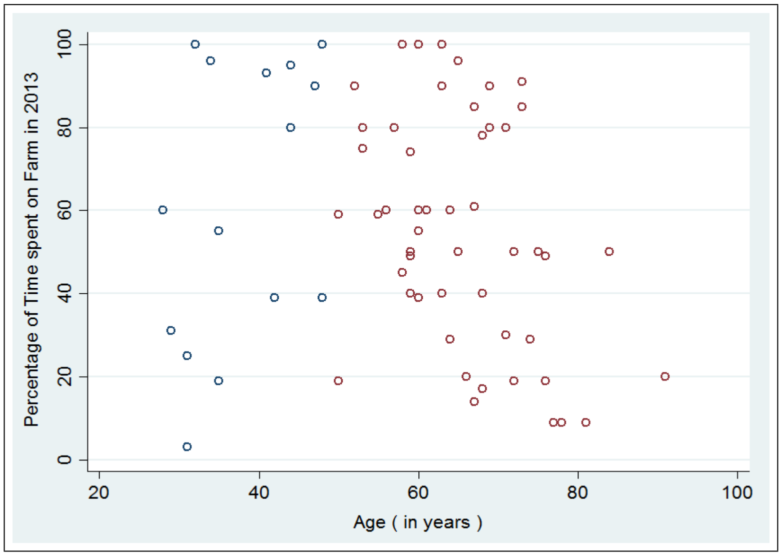



Figure 2.3 Time spent farming versus farmer age

Figure 2.3 plots labor hours devoted to farming (vertical axes) against farmer age (horizontal axes). Respondents below 50 years of age are shown in blue and respondents above 50 years in age are shown in red. The simple correlation between the labor time spent farming with farmer's age is -0.368 (p-value 0.008) for farmers 50 years of age or older and 0.4301 (p-value 0.109) for farmers younger than 50 years. The overall correlation is -0.196 (p-value 0.115).

### 2.3.2 Management practices

Table 2.2 reports the respondents' sources for nutrient management advice and the field types for which it is sought. The main source of nutrient management advice is from a professional agronomist at the cooperative: 92.4% reported receiving nutrient management advice from this source. The proclivity to seek advice from the cooperative agronomists is not surprising considering the survey respondents in our sample are producer members of the cooperative. One-third of respondents sought nutrient management advice from more than one source. Among the 72 respondents who responded to advice-source questions, 55.6% received field-

[5]Farm size greater than 2000 acres have been excluded to make the figure presentable.

Table 2.2 Nutrient management decisions and advice. **Shared decision making** denotes the percentage of respondents who shared responsibility with others for the nutrient management decisions. **Source of advice** denotes the percentage of respondents who reported they received advice from the four sources indicated. **Advice regarding** denotes the percentage of respondents who received advice pertaining to the three field types indicated.

| | | | |
|---|---|---|---|
| **Shared decision making** | | | |
| | N | Yes (%) | No (%) |
| | 77 | 36.36 | 63.64 |
| **Source of advice** | | | |
| | N | Yes (%) | No (%) |
| ISU Extension | 66 | 21.21 | 78.79 |
| Agronomist at Farming Co-op | 66 | 92.42 | 7.58 |
| Agronomist at Professional Consulting Firm | 66 | 9.09 | 90.91 |
| Other Farmers | 66 | 19.7 | 80.3 |
| **Advice regarding** | | | |
| | N | Yes (%) | No (%) |
| Best Producing field | 72 | 52.78 | 47.22 |
| Average Producing field | 34 | 55.88 | 44.12 |
| Under Performing field | 38 | 52.63 | 47.37 |

specific nitrogen management advice on at least one of their fields, and 51.39% of respondents received nitrogen management advice for at least two of the fields they manage.

In informal discussions with agronomists at the cooperative and also Iowa State University, agronomists mentioned that producers may be following older recommendations and 'rules of thumb' for nitrogen that are no longer used, and despite their efforts to change those recommendations, producers still use them. In table 3 the producers' responses to how they use the advice from the sources they reported in table 2, are reported. Forty of the respondents

Table 2.3 Impact of Nitrogen Management Advice. Table reports response to question on whether received advice was followed.

| | N | % |
|---|---|---|
| I followed the advice exactly | 13 | 32.50 |
| Based on the advice, I made big adjustment | 8 | 20.00 |
| Based on the advice, I made small adjustment | 18 | 45.00 |
| I did not follow the advice at all | 1 | 2.50 |
| Total | 40 | 100 |

answered this question, and of those, about one-third said they followed the advice exactly, and nearly all of the remaining respondents said they used the advise to adjust their plans.

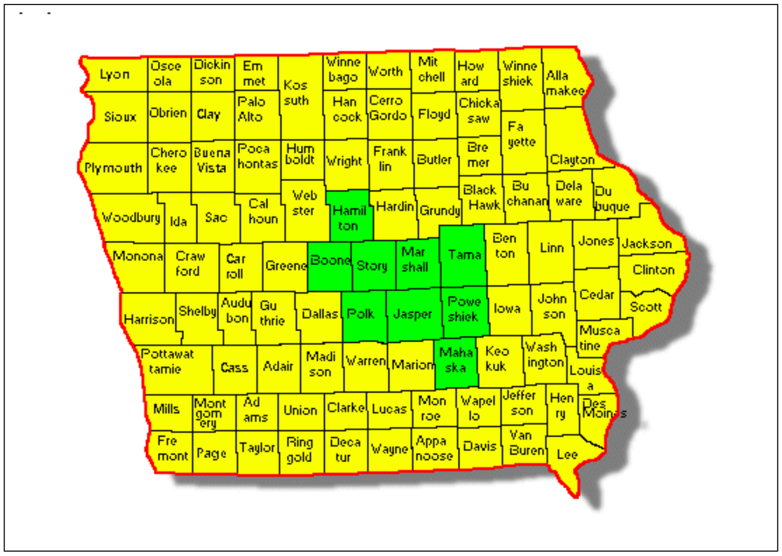



Figure 2.4 Survey respondents geographical location

Figure 2.4 illustrates the Iowa counties in which the respondents' fields are located. Sampled fields are concentrated in Jasper county (41.46%) and Story county (24.68%). The remaining fields (33.77%) are located in Boone, Hamilton, Mahaska, Marshall, Polk, Poweshiek and Tama counties.

Table 2.4 reports summary statistics for field size, soil quality, as measured by the proportion of highly erodible land (HEL), and soil nitrogen requirements.

Sample mean CSR rating on the best producing, average producing, and under performing fields is 82.8, 74.41, and 62.14, respectively. HEL shows a similar pattern with the mean and median HEL on best producing fields at 20.23% and 0%, respectively. The mean and median percentage values on average performing fields increase to 45.55% and 23.8%, respectively. The mean and median values for HEL on under performing fields is 53.1% and 61.6%, respectively.

The percentage of HEL land is negatively correlated with field CSR. The correlation coefficient between field CSR and the percentage of HEL land is -0.287 (p-value 0.016) for best producing field, -0.200 (p-value 0.273) for average producing field and -0.676 (p-value $< 0.000$)

Table 2.4 Field characteristics.

| **Field Size (acres)** | | | | | | |
|---|---|---|---|---|---|---|
| Field type | N | Mean | Median | Std. Dev. | Min. | Max. |
| Best producing | 77 | 85.84 | 77 | 65.27 | 9 | 350 |
| Average producing | 34 | 85.56 | 70 | 58.31 | 10 | 300 |
| Under performing | 38 | 49 | 40 | 33.50 | 5 | 160 |
| **CSR rating** | | | | | | |
| Field type | N | Mean | Median | Std. Dev. | Min. | Max. |
| Best producing | 70 | 82.8 | 84.5 | 7.68 | 60 | 95 |
| Average producing | 32 | 74.41 | 73.5 | 5.91 | 65 | 85 |
| Under performing | 35 | 62.14 | 60 | 15.62 | 30 | 90 |
| **Highly erodible land (% of total acres)** | | | | | | |
| Field type | N | Mean | Median | Std. Dev. | Min. | Max. |
| Best producing | 77 | 20.23 | 0 | 33.81 | 0 | 100 |
| Average producing | 34 | 45.55 | 23.81 | 45.49 | 0 | 100 |
| Under performing | 38 | 53.06 | 61.25 | 45.31 | 0 | 100 |
| **Estimate of nitrogen required (in lbs./acre)** | | | | | | |
| Field type | N | Mean | Median | Std. Dev. | Min. | Max. |
| Best producing | 76 | 167.84 | 162.5 | 71.11 | 0 | 500 |
| Average producing | 34 | 133.82 | 150 | 98.94 | 0 | 530 |
| Under performing | 38 | 135.34 | 150 | 73.56 | 0 | 300 |

for under performing field. This negative correlation is more pronounced on under-performing fields.

Respondents were asked to report nitrogen requirements on their fields, given the yields that they expected. Respondents then were asked a follow up question that asked how confident they were about their field's nitrogen requirements. Respondents chose between pre-specified confidence intervals intended to provide the most accurate description of subjective confidence. Options included a 95% level of confidence, a 90% level of confidence, and so on (options of 80% and 50% confidence intervals were included).[6] Finally, respondents were allowed a *Not*

[6]The interval options provided upper and lower values for the lbs./acre of nitrogen associated with each percentage-based confidence interval.

Table 2.5 Farmers' Confidence Regarding Nitrogen Requirements. Table reports respondents confidence regarding nitrogen requirements on managed fields.

| Field type: | Best producing | Average performing | Under performing |
|---|---|---|---|
| 95% confident | 20 (27.40%) | 7 (24.14%) | 9 (28.13%) |
| 90% confident | 18 (24.66) | 4 (13.79) | 7 (21.88) |
| 80% confident | 14 (19.18) | 8 (27.59) | 9 (28.13) |
| 50% confident | 18 (24.66) | 9 (31.03) | 5 (15.63) |
| Not sure | 3 (4.11) | 1 (3.45) | 2 (6.25) |
| Total | 73 | 29 | 32 |

*Sure* option to reflect the case of little or no confidence in beliefs about the fields' nitrogen requirements. Table 2.5 summarizes the responses to this question.

### 2.3.3 Subjective yield and rainfall expectations

Table 2.6 Summary Statistics: Expected Corn Yield (bu./acre).

| Field type | N | Mean | Median | Std. Dev. | Min. | Max. |
|---|---|---|---|---|---|---|
| Best producing | 68 | 202.06 | 200 | 21.20 | 150 | 250 |
| Average producing | 26 | 181.32 | 185 | 16.96 | 140 | 220 |
| Under performing | 29 | 172.32 | 175 | 19.71 | 125 | 210 |

Table 2.6 provides summary statistics about expected corn yields in bushels per acre by field type. The sample mean and median values for expected yields vary across field types as expected. Recall that the survey instrument allowed the respondent to select a field that they manage based only on the criteria, that is "best producing". It is no surprise that expected yields are highest on fields identified by respondents as "best producing" than on the other two field types.

Figure 2.5 plots the quantity of nitrogen applied (measured as cumulative sum of the monthly nitrogen applications the respondent reported) against the elicited subjective expected corn yields. The correlation between elicited expected corn yield and nitrogen applied for the *Best producing fields* is 0.355 (p-value 0.003). The positive correlation suggests that farmers who apply more nitrogen expect higher yields on their *Best producing fields*.

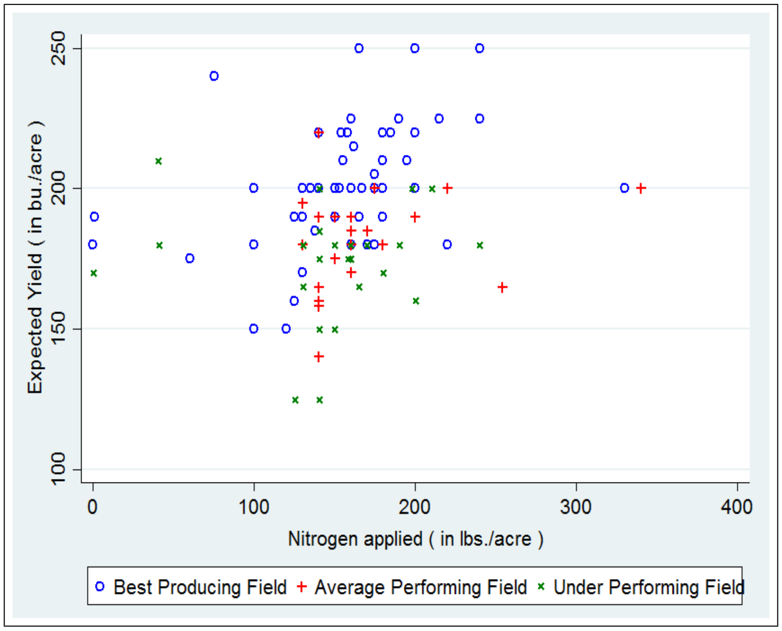



Figure 2.5 Nitrogen Application and Expected Yield

Figure 2.6 plots field CSR (horizontal axes) against subjective expected yield. The correlation between CSR and expected yield for *Best producing fields* is 0.313 (p-value 0.009).

### Subjective yield distributions

The producers' responses allowed capturing of a total of 121 subjective corn yield distributions using the SPIES approach. Among the 68 "best producing" field yield distribution measurements, 3 violations of monotonicity of cumulative probability were found. Of the 53 under-performing field measurements, 1 violation of monotonicity of cumulative probability was found. It should be noted that these violations occurred despite being prompted by the survey software of the problem.

A second axiom of probability checked is whether the probabilities of corn yield realizations sum to one. Note that unlike the monotonicity axiom, no warning message was provided to respondents if elicited probabilities, more precisely *chances out of 100* violated the axiom. Among the 68 best producing field measurements 5 did not satisfy the adding up axiom. Among the 53 average producing or under-performing measurements 2 violated the adding up axiom.

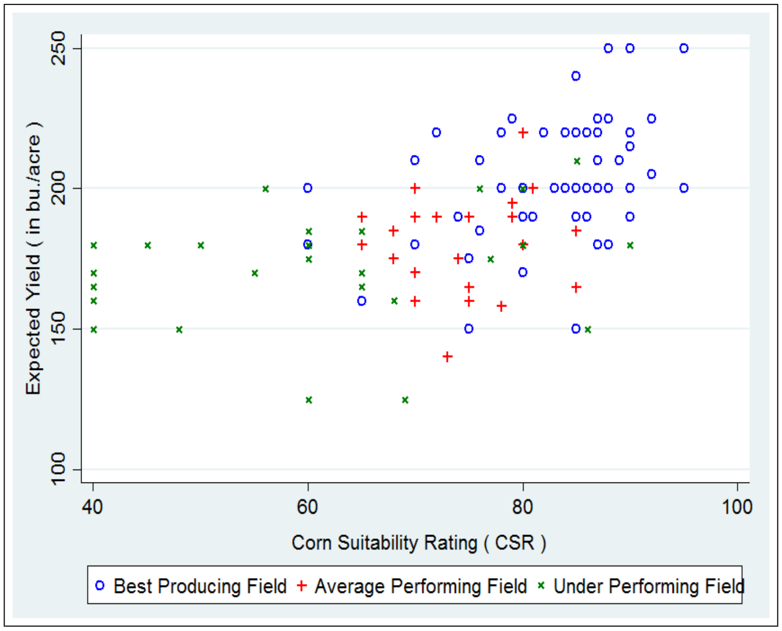



Figure 2.6 Corn Suitability Ratings (CSR) and Expected Yield. Blue circle represent the response for *Best producing fields*, red plus indicates *Average producing fields* and, *Under performing fields* are indicated by green cross.

The SPIES approach was used to measure 73 subjective rainfall distributions for the *Best producing fields*. Violations of probability axioms were more prevalent than with yield distribution measurements. 64 of the 73 cases satisfies the monotonicity of cumulative probability axiom. Among the 9 violations, 6 placed very high values (as high as 96 chances out of 100) on drought conditions as indicated by rainfall totals less than 0.5 inches. The remaining 3 violations appear to be cases of epistemic uncertainty. A check of the adding up axiom indicated 29 violations.

## 2.4 Discussion and Suggestions for Future Survey Designs

This section discusses what is learned from the pilot survey design and implementation including shortcomings, and contemplates how future survey-based research in similar agricultural settings may be conducted.

### Response rate

Although, an approximate response rate of 19% is reasonable, it was below the expected response rate. The relatively low response rate is explained by the timing of the survey and also the lack of follow-up to encourage greater participation. Survey invitations were sent and responses collected in June. In Iowa, planting is nearly complete at this time but additional field work like spraying keeps producers very busy, and this can be a busy time for producers with livestock as well. This may have contributed significantly to the response rate. Further, the late survey necessarily means the responses capture more information about what actually happened instead of what was planned and anticipated. Therefore, the ideal time for this survey and the stated goals is perhaps February or early March when nutrient management decisions are formulated and much uncertainty still exists about weather and planting conditions.

### Potential Ambiguity in Questions

Producers were asked to report about the crop growing in 2014. A few respondents reported that the current crop was soybeans, yet, their reported yields were almost certainly for corn on the identified field. The prior questions in the survey may have caused this. The questions prior to the expected yield question were related to nitrogen usage on the field, and it was assumed that respondents were thinking about a prior year in which the field was planted to corn since nitrogen application is not relevant to soybean. This effectively makes their responses hypothetical ones. This ambiguity needs to be addressed in future iterations of this survey.

N-P-K ratio is very standard labeling of fertilizer composition that summarizes the proportion of Nitrogen, Phosphorus and Potassium. Focus group meetings indicated the N-P-K rating convention is familiar and well-understood by farmers. An example was provided in the survey. In spite of this, several farmers have reported numbers inconsistent with percentages. It is likely that they have reported quantities rather than percentages. This should be clarified in future studies.

There is a negligible proportion of farmers in the sample who used manure. Questions pertaining to manure was included in the survey as suggested in the focus group. Because of the lower prevalence of livestock production in Central Iowa than, for example, North Western or Western Iowa, the population of producers are perhaps less likely to use manure and depend more on commercial nitrogen.

It is assumed that the expected yield elicited from respondents is the mean of the expected yield distribution. However, this is not guaranteed, nor is it guaranteed that a producer reported, for the relevant questions, a mean in all questions rather than some other measure of central tendency, like median. It may be assumed that if the respondent reports a mean (and not median) for one of the field, they are highly likely to report on the mean for the other field, too. The cumulative probability response was used to check this. If the respondent elicited a chance greater than or equal to 50 that either the expected yield will be less than 90% of Expected Yield threshold or greater than 110% of Expected Yield threshold or both, it is certain that they have not reported a median. If a median was elicited as expected yield, less than or greater than cumulative probability around expected yield could not have been greater than or equal to 50. 43% of the farmers who grow corn on at least one of the fields, have reported a value that corresponds to mean. For the rest it is still not known and have been assumed that the reported Expected yield is the mean.

There was a programming error in the cumulative probability of rainfall for the *Average Producing* or *Under Performing Field*. Hence, farmers' rainfall probability are not reported. Moreover, since county locations of both the fields are same for most of the respondents, disparate rainfall beliefs in the two fields are not expected. It is seen that the inconsistent responses in the rainfall probability is relatively higher. One of the reasons could be the external anchoring aspect of it in spite of the fact that the actual July rainfall for 2014 was indeed 4.5 inches, which is close to the average July rainfall of 4.3 inches. Also, the chances of rainfall as low as 2 inches is a drought situation which is not very common. But since farmers have encountered severe drought in 2012, high chances of a drought could be a representation bias[7]. Similarly, a situation of rainfall as low as 0.5 inches or below has never occurred before and can be objectively considered to have a probability of 0. But it is seen that most of

[7]Refer to Tversky and Kahneman (1974) for representation bias

the respondents have placed positive weight on this event. This is an anchoring bias[8] where respondents believe that if they have been asked to state the probability of occurrence of an event, then there must be positive chances of occurrence of the event.

**Anchoring effects**

In elicitation of the subjective probabilities, it was learned that self-anchoring as in the expected yields is a better way to elicit responses as it reduces representative bias and anchoring bias which was observed in the elicitation of subjective rainfall probabilities. It also came to notice during the survey that the rainfall probability elicitation for the second field was jumbled due to a programming error, as respondents were asked to report less than subjective probabilities for 6.5 and 8.5 inches of rainfall instead of greater than probability. Since this was the last of probability elicitation, many respondents have assumed it was an error and reported what was intended to be asked but many responded to the words framed in the question. Although this was undesirable, it did not reflect the study significantly as it had been correctly framed for the first field. Since for most farmers two fields have been in the same county, subjective weather beliefs for first field are representative for second field too (unless farmers are biased in their beliefs about weather across fields). Nonetheless, this unintended misplacement of words have provided a hint that randomization of questions across respondents for the different thresholds of yield or rainfall may make the responses robust minimizing its dependence on format used.

## 2.5 Conclusion

This chapter presents a overview of a pilot study of U.S. midwest farmers decision making that was conducted in Iowa in 2014. A rationale is provided for the methods used. Summary statistics for some survey findings are reported and shortcomings of the pilot study are discussed and lessons for future research on subjective belief measurement in agriculture are learned. The pilot study survey instrument is reported in an appendix.

[8]Refer to Tversky and Kahneman (1974) for anchoring bias

The main conclusion that can be drawn is that survey methods that employ direct subjective belief measurement appear to be a viable approach for studying decision making processes and belief formation in uncertain production environments.

## 2.6 Appendix: The Survey Instrument

Please enter your access code on the top left hand side of the survey invitation letter you received from *XXX* Cooperative. [ ]

Please re-enter the access code and confirm [ ]

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

**Introduction**

Thank you for agreeing to participate in our survey. The goal of this project is to learn more about the decision processes used by producers when choosing how, when, and how much fertilizer to apply to their fields. The data we collect will be used for this purpose alone.

Any results we report will be sufficiently aggregated to mask individual responses. Your personal responses and your identity will be kept strictly confidential. You can view the Confidentiality Form Here.

The survey should take at most 25-30 minutes to complete. Upon completion, we will send you a formal signed confidentiality agreement that describes the safeguards that will be used to protect your data and a $50 check. If you are interested, we will also share a copy of our study results when they are ready for distribution.

Participating in this survey is an opportunity to help advance research in the broad area of decision making under uncertainty. Your participation will provide new insights into the decision processes of farmers. Please think carefully about each response you provide and remember that the only wrong answer to a question is one that is disingenuous.

We will begin the survey by asking general questions about your farming business. Questions about fertilizer application decisions on specific fields that you farm will follow.

If you have any questions please contact a project leader:

| | |
|---|---|
| Keri Jacobs | Quinn Weninger, |
| Department of Economics | Department of Economics |
| Iowa State University | Iowa State University |
| email: kljacobs@iastate.edu | email: weninger@iastate.edu |
| Phone: (515) 294-6780 | Phone (515) 294-8976 |

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

**General Information I**

We would like to know the number of acres on which you are responsible for making nutrient management decisions and, in particular nitrogen application decisions.

Acres that I own personally ☐ acres

Acres that I lease from others ☐ acres

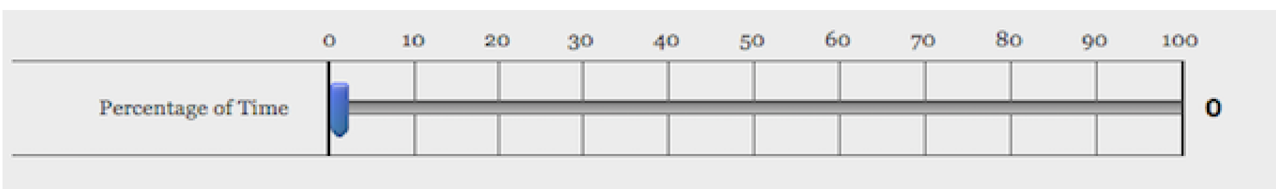
Acres that do not fit the categories above, e.g. absentee landowner ☐ acres

Do you share nutrient management decision responsibilities on these acres with other individuals, e.g., a partner who is or is not a family member? If so how many people do you consult?

○ Yes ☐

○ No

Please estimate the percentage of your total working hours in 2013 that were spent on activities related to your farm business?




Please specify the count of each kind of animal you have on your farm.

☐ Slaughter or Feeder Cattle ☐

☐ Immature Dairy Cattle ☐

☐ Mature Dairy Cattle ☐

☐ Swine 55 pounds or more ☐

☐ Swine under 55 pounds ☐

☐ Sheep and Lambs ☐

☐ Horses ☐

☐ Turkeys over 7 pounds

☐ Turkeys under 7 pounds

☐ Broiler or layer chickens 3 pounds or more

☐ Broiler or layer chickens under 3 pounds

Do you submit or annually update a Manure Management Plan to the Iowa Department of Natural Resources?

○ Yes

○ No

Did you buy / sell or are you planning to buy / sell manure in 2014? If you answer yes specify the amount bought or sold.

☐ Yes I will buy or have bought (in pounds)

☐ Yes I will sell or have sold (in pounds)

☐ No

How many distinct fields do you manage? fields

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

**Fertilizer management**

Next we will ask a series of questions about management practices on your best producing field. Please keep this particular field in mind when you answer the following questions.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*[Note: Questions in this section apply to the respondents Best Producing field]*

In what county is your best field located?

*[Note: Respondent is presented with drop down list of Iowa counties ]*

What is your best fields corn suitability rating (CSR)?

CSR =

How many of your best fields total acres are classified as highly erodible land (HEL)?

☐ acres

How large is your best field?

☐ acres

How would you rate the fertility on your best field?

○ Poor

○ Less than adequate

○ Adequate

○ Better than adequate

○ Great

What crop was planted on your best field in the 2012 growing season? (If more than one crop was planted indicate by clicking multiple buttons.)

☐ Corn

☐ Soybean

☐ Other ☐

What crop was planted on your best field in the 2013 growing season? (If more than one crop was planted indicate by clicking multiple buttons.)

☐ Corn

☐ Soybean

☐ Other ☐

What crop(s) did you or will you plant on your best field in 2014? (If more than one crop will be planted indicate by clicking multiple buttons.)

☐ Corn

☐ Soybean

☐ Other ☐

What tillage practice did you use or do you plan to use on your best field during the 2014 growing season?

☐ Conventional Tillage

☐ Minimum Tillage

☐ Mulch Tillage

☐ No Till

☐ Conservation Tillage

☐ Strip Tillage

☐ High residue

Have you used or are you planning to use a nitrogen test on your best field?

☐ No

☐ Yes, I have done or plan to do a soil test

☐ Yes, I have done or plan to do a plant tissue test

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

[ *Note: This page appears only if the respondent has chosen "Yes, I have done or plan to do a soil test" and/or "Yes, I have done or plan to do a plant tissue test"* ]

In what year and month was the most recent Nitrogen test conducted on your best field?

☐ Year ☐ Month

Are you planning additional Nitrogen tests this growing season on your best field?

☐ No

☐ Yes, I plan to test for Nitrogen before planting

☐ Yes, I plan to test for Nitrogen after planting

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

We next want you to think about and estimate the Nitrogen content in your your best fields soil today. We realize it may be difficult to know the exact nitrogen concentration. Please make your best estimate.

Based on your best estimate of the current nitrogen concentration on your best field, how many pounds per acre do you think are needed to achieve your expected yield?

☐ pounds per acre

*[Note: Denote the response to this questions as* ***Z0*** *lbs./acre]*

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Which of the following statements best describes your confidence in your estimate of your best fields Nitrogen needs?

○ I'm confident the field needs more than *[**97.5% of Z0**]* pounds/acre but less than *[**102.5% of Z0**]* pounds/acre.

○ I'm confident the field needs more than *[**95% of Z0**]* pounds/acre but less than *[**105% of Z0**]* pounds/acre.

○ I'm confident the field needs more than *[**90% of Z0**]* pounds/acre but less than *[**110% of Z0**]* pounds/acre.

○ I'm confident the field needs more than *[**75% of Z0**]* pounds/acre but less than *[**125% of Z0**]* pounds/acre.

○ I'm not sure at all about the fields' Nitrogen needs.

We would now like to ask about Nitrogen applications. Please select the months when commercial nitrogen was/will be applied to your best field.

☐ September, 2013

☐ October, 2013

☐ November, 2013

☐ December, 2013

☐ January, 2014

☐ February, 2014

☐ March, 2014

☐ April, 2014

☐ May, 2014

☐ June, 2014

☐ July, 2014

☐ August, 2014

[ *Note: Denote **NM3** as the latest month that nitrogen was applied.* ]

Please select the months when manure was/will be applied to your best field.

☐ September, 2013

☐ October, 2013

☐ November, 2013

☐ December, 2013

☐ January, 2014

☐ February, 2014

☐ March, 2014

☐ April, 2014

☐ May, 2014

☐ June, 2014

☐ July, 2014

☐ August, 2014

[ *Note: Denote **MM1** as the last month that manure was applied.* ]

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

In thinking of your commercial fertilizer use, please indicate the percentages of N/P/K given in the reported guaranteed analysis of the mixed grades or straight materials. For example, a common fertilizer might carry a guaranteed analysis of 10% nitrogen fertilizer, 34% phosphorus, and 0% potassium and be reported as a 10-34-0 mixed grade.

What fertilizer did (will) you apply in [***NM1***] on your best field?

☐ N ☐ P ☐ K

What fertilizer application method was (will be) used in [***NM1***] on your best field?

☐ Broadcast

☐ Anhydrous

☐ Side banding

☐ Seed Furrow

☐ Side dress

☐ Late side dress

☐ Others (Please Specify) ☐

How many pounds of nitrogen were applied on your best field in [***NM1***]?

☐ pounds per acre

[ *Note: Let **QN1** denote lbs./acre applied in month **NM1**.* ]

What fertilizer did (will) you apply in [***NM2***] on your best field?

☐ N ☐ P ☐ K

What fertilizer application method was (will be) used in [***NM2***] on your best field?

☐ Broadcast

☐ Anhydrous

☐ Side banding

☐ Seed Furrow

☐ Side dress

☐ Late side dress

☐ Others (Please Specify) ☐

How many pounds of nitrogen were applied on your best field in [***NM2***]?

☐ pounds per acre

[ *Note: Let **QN2** denote lbs./acre of nitrogen applied in month **NM2**.* ]

What fertilizer did (will) you apply in [***NM3***] on your best field?

☐ N ☐ P ☐ K

What fertilizer application method was (will be) used in [***NM3***] on your best field?

☐ Broadcast

☐ Anhydrous

☐ Side banding

☐ Seed Furrow

☐ Side dress

☐ Late side dress

☐ Others (Please Specify) ☐

How many pounds of nitrogen were applied on your best field in [***NM3***]?

☐ pounds per acre

[ *Note: Let **QN3** denote lbs./acre of nitrogen applied in month **NM3***]

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

Did you do a manure Nutrient Content Analysis prior to applying manure on your best field in [***MM1***]?

○ Yes

○ No

What type and quantity of manure did you apply on your best field in [*MM1*]? pounds/acre (gallons/acre if applied in liquid form)

☐ Beef

☐ Dairy

☐ Swine

☐ Poultry

☐ Others

Which manure application methods were used on your best field in [*MM1*]?

☐ Manure Spreader

☐ Tanks

☐ Injectors

☐ Pumps

☐ Umbilical system

☐ Other (please specify)

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

**Weather and Yield**

We would now like to ask questions about the weather conditions and yield you expect in the upcoming 2014 growing season.

Keeping in mind the spring weather conditions , which of the following is closest to the date you have planted or will most likely plant your crop?

☐ April, 1, 2014

☐ April, 15, 2014

☐ May, 1, 2014

☐ May, 15, 2014

☐ June, 1, 2014

☐ June, 15, 2014

☐ July, 1, 2014

☐ July, 15, 2014

☐ Not sure at all about when I will have the crop planted

If planting is completed on the date you expect, during which of the following periods is pollination most likely to occur on your best field?

☐ May 1 - May, 15, 2014

☐ May 15, - May 31, 2014

☐ June 1 - June, 15, 2014

☐ June 15, - June 31, 2014

☐ July 1 - July, 15, 2014

☐ July 15, - July 31, 2014

☐ August 1 - August, 15, 2014

☐ August 15, - August 31, 2014

☐ Not sure at all about the pollination period

If you follow the nitrogen application schedule described earlier, how many bushels do you expect your best field will yield?

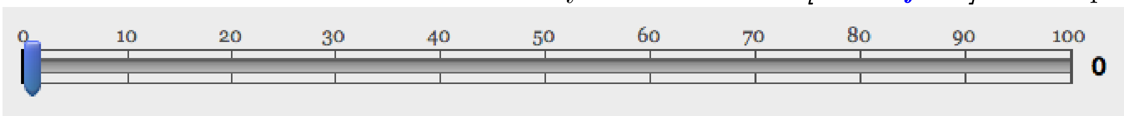
☐ Bushels per acre

[ *Note: Denote the response to the above question as [**Y1**] Bu./acre.* ]

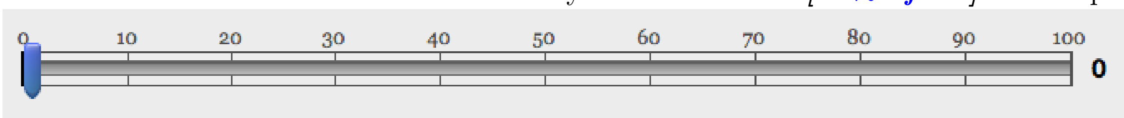
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

You indicated that following the current nitrogen plan, you expect a yield of *[**Y1**]* bushels per acre on your best field.Taking this into account please answer the following questions.

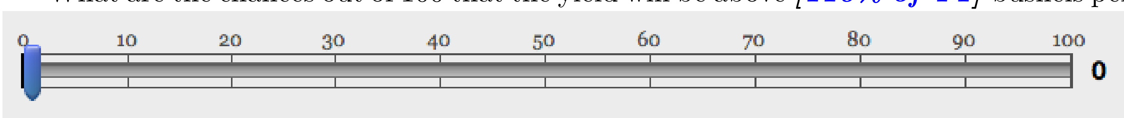
What are the chances out of 100 that the yield will be below *[**90% of Y1**]* bushels per acre.



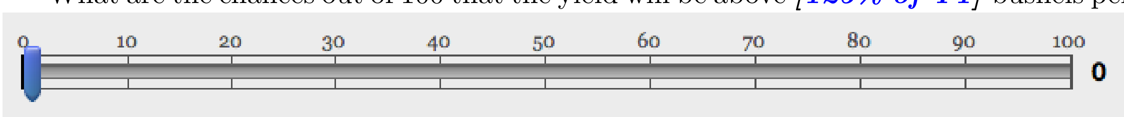

What are the chances out of 100 that the yield will be below *[**75% of Y1**]* bushels per acre.




What are the chances out of 100 that the yield will be above *[**110% of Y1**]* bushels per acre.




What are the chances out of 100 that the yield will be above *[**125% of Y1**]* bushels per acre.




[ *Note: If the response violates monotonicity of cumulative probability, the following error message appears; "Your response indicates that the chances average yield falls below [75% of Y1] bushels per acre are greater than the chances average yield falls below [90% of Y1] bushels per acre. Please review and confirm your responses to the last two questions, or "Your response indicates that the chances average yield exceeds [110% of Y1] bushels per acre are less than the chances average yield exceeds [125% of Y1] bushels per acre. Please review and confirm your responses to the last two questions"* ]

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

We would now like you to describe how yield might change if you were to apply different amounts of nitrogen fertilizer. You indicated that you plan to apply [***QN3***] pounds/acre in your [***NM3***] application.

Suppose instead you applied [***115% of QN3***] pounds per acre at the [***NM3***] application. How many bushels per acre would your best field now yield?

☐ Bushels per acre

Suppose instead you applied [***130% of QN3***] pounds per acre at the [***NM3***] application. How many bushels per acre would your best field now yield?

☐ Bushels per acre

Suppose instead you applied [***85% of QN3***] pounds per acre at the [***NM3***] application. How many bushels per acre would your best field now yield?

☐ Bushels per acre

Suppose instead you applied [***75% of QN3***] pounds per acre at the [***NM3***] application. How many bushels per acre would your best field now yield?

☐ Bushels per acre

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

**Weather Conditions and Yield**

Data from the National Oceanic and Atmospheric Administration indicates that the historical average rainfall in central Iowa during the month of July is 4.3 inches per month.

Suppose 6.5 inches of rain falls during the month of July. How many bushels per acre would your best field now yield?

☐ Bushels per acre

Suppose 8.5 inches of rain falls during the month of July. How many bushels per acre would your best field now yield?

☐ Bushels per acre

Suppose 2 inches of rain falls during the month of July. How many bushels per acre would your best field now yield?

☐ Bushels per acre

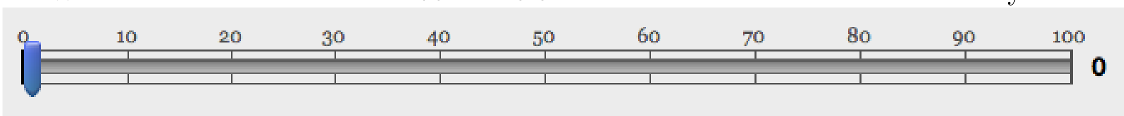
Suppose 0.5 inches of rain falls during the month of July. How many bushels per acre would your best field now yield?

☐ Bushels per acre

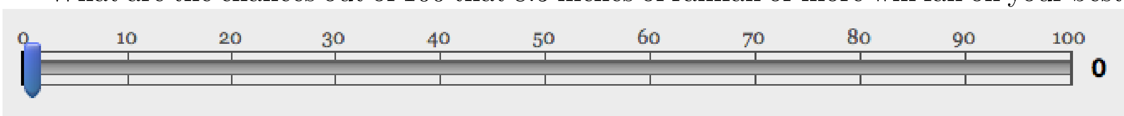
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

We realize that rainfall can be difficult to predict. We would like you to estimate rain that will fall on your best field during July of 2014. Keep in mind that the historical average monthly rainfall in July is 4.3 inches.

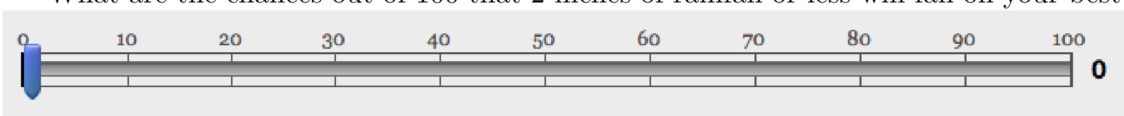
What are the chances out of 100 that 6.5 inches of rainfall or more will fall on your best field?



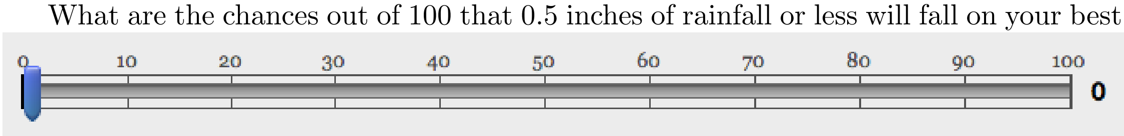

What are the chances out of 100 that 8.5 inches of rainfall or more will fall on your best field?




What are the chances out of 100 that 2 inches of rainfall or less will fall on your best field?




What are the chances out of 100 that 0.5 inches of rainfall or less will fall on your best field?




[ *If the response violates the monotonicity of cumulative probability, the following error message appears; "Your response indicates that the chances average rainfall in July 2014 being more than 8.5 inches is greater than the chances average rainfall in July 2014 being more than 6.5 inches. Please review and confirm your responses to the last two questions" or "Your response indicates that the chances average rainfall in July 2014 being less than 2 inches is lesser than the chances average rainfall in July 2014 being less than 0.5 inches. Please review and confirm your responses to the last two questions"* ]

[ *Note: This marks the end of section specific to the Best producing field* ]

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

*[For each respondent an* ***average producing*** *or* ***under performing*** *field was chosen at random for questions pertaining to this section. Let the randomly chosen field be [****FTYP****] ]*

Next we will ask a series of questions about management practices on one of your [***FTYP***] performing fields. Please keep this particular field in mind when you answer the following questions.

⊙ Press Next to Continue

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

[ *Note: The questions in this section are directed at the respondent's [**FTYP**] performing fields.* ]

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

We will close with few more general questions.

Have you received Nitrogen management advice from any of the following sources in the past year?

☐ Iowa State University Extension Services

☐ Agronomists at my Farming Coop

☐ Agronomists at a professional consulting firm

☐ Other farmers

Have you ever received a specific Nitrogen management recommendation for your best field?

○ Yes

○ No

Have you ever received a specific Nitrogen management recommendation for your [***FTYP***] producing field?

○ Yes

○ No

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

[ *Note: The next questions are asked only if the respondent received nitrogen management recommendations for their best or [FTYP] performing field or both.* ]

How much influence did the advice you received affect the nitrogen management plan that you eventually followed?

○ I followed the advice exactly

○ Based on the advice, I made big adjustments to my Nitrogen management strategy

○ Based on the advice, I made small adjustments to my Nitrogen management strategy

○ I did not follow the advice at all

We will conclude our survey with a few questions about yourself and your farming experience.

How old are you?

☐

How long have you been farming?

☐

What is the highest level of schooling that you completed?

○ High school

○ Four year undergraduate degree

○ Graduate degree

Thank you for participating in this study!

☐ Yes, please send me a copy of the study findings !

# CHAPTER 3. NITROGEN AND WEATHER CONDITIONAL YIELD DENSITY: SOME EMPIRICAL FACTS

## 3.1 Introduction

In this chapter, input conditional yield distribution is modeled, which highlights the effect of nitrogen, weather and their interaction on the moments of the yield density. While said so, it is important to note that this chapter is a precedent to the main chapter, chapter 4, where insights from this chapter have been used to build up the main results of the dissertation. This chapter highlights the role of nitrogen and weather inputs, and their interaction in a stochastic yield production function. The purpose of this chapter is to develop an objective model of yield nitrogen mapping that can be used in the next chapter to provide a comparable benchmark for the corresponding subjective belief that has been measured using the survey instrument described in chapter 2. The modeled yield distribution provide estimates of moments of the yield distribution as a function of nitrogen and weather variables. This chapter also discusses the economic implications of the estimated moments of the yield distribution. Researchers have used several alternative techniques to model the yield distribution. The method used in this chapter provides a flexible modeling framework and produces results that are comparable with the results of previous research studies. Moreover, a discussion of the yield nitrogen relationship and its relevance for a nitrogen-decision making for a farmer is provided.

Due to the stochastic nature of agricultural production function, distribution of agricultural output (yield) is at the heart of agricultural decision making. Knowledge of yield distribution is fundamental to agricultural economics. Some examples where yield distribution plays a significant role includes provision and pricing of agricultural financial products, technological developments and performance of existing technologies, analyzing producer's input choice and technology adoption decisions, agriculture linked environment conservation policies etc. Therefore, modeling of crop yield density and response of crop yield to factor inputs have received attention from agronomists and agricultural economists. Not only the mean but the higher moments of the yield distribution including variance and skewness are of significant interest to researchers. Although, economists have used variance as a measure of risk (Rothschild and Stiglitz, 1970; Just and Pope, 1978), agricultural economist have paid significant attention to skewness too as a measure of risk (Day, 1965; Antle, 1983; Babcock and Hennessy, 1996; Du et al., 2012). Yield skewness measures the asymmetry of the yield distribution. It compares the cumulative probability in the left tail versus the right tail of the distribution around some central value (mostly mean) of the yield distribution.

The chapter progresses by providing a literature review of some popular distributions that have been used to model the crop yield density in section 3.2. Section 3.3 outlines the method of generalized linear model (GLM) with beta distribution and discusses its use in context to the modeling of yield distribution. Section 3.4 describes the nitrogen trial experimental data that has been used for the empirical analysis. Section 3.5 discusses the findings of the chapter by outlining the effect of the weather and nitrogen on the moments of the yield distribution and their economic implication. Section 3.6 concludes. An appendix to the chapter is provided in the section 3.7.

## 3.2 Literature Review

Researchers have employed several techniques to model the crop yield distribution. Research based on aggregate time series data, have generally fitted unconditional distributions to detrended yields (Moss and Shonkwiler, 1993; Ramirez, 1997; Just and Weninger, 1999; Ramirez et al., 2003; Harri et al., 2009; Zhu et al., 2011); aggregate yields have been conditioned on weather if a conditional distribution is fit (Gallagher, 1986; Thompson, 1986, 1988; Kaylen and Koroma, 1991; Chen et al., 2004; Schlenker and Roberts, 2006, 2009; Tack et al., 2012). Modeling conditional yield distribution has been more popular among researchers who have used farm level data. However, as suggested by some researchers (Just and Weninger, 1999; Claassen and Just, 2011), aggregate county or state level time series yield data is not representative of individual farm level yield data and hence, must be used with caution. Recently, Harri et al. (2009) provided a reconciliation of earlier studies that have modeled yield density. Using county level yield for different crops across the U.S., the study showed that the disparate results across different studies regarding the crop yield density are actually not contradictory. They find support in favor of findings from previous studies, based on which they suggest that yield distributions are localized from crop to crop and therefore, results from a particular crop and region cannot be universally applied to some other crop or region.

The Von Liebig production technology or the Linear Plateau function (LRP) is often used by agronomists and agricultural economists to approximate the yield response to nitrogen (Cerrato and Blackmer, 1990; Babcock and Blackmer, 1992). Recently, researchers have extended the original LRP function to stochastic LRP using random parameters model which is reported to be a better fit than its non-stochastic counterpart (Tembo et al., 2008; Tumusiime et al., 2011; Boyer et al., 2013).

Alternatively, smooth production functions (using a functional form for the production function, which is mathematically differentiable compared to LRP, which has a kink) have been used by Just and Pope (1979); Antle (1983); Babcock and Hennessy (1996); Du et al. (2012). Just and Pope (1978, 1979) was an econometric innovation in the use of smooth stochastic production function (J-P production function henceforth) as it untied the effect of input on the mean yield from the effect of input on the variability of the output. In spite of the flexibility of J-P production function in terms of the yield variability, Antle (1983) showed that it imposes arbitrary restrictions on various input elasticities. An alternative flexible moment based framework for estimation of a stochastic production function was introduced by Antle (1983) (which has been relatively less popular compared to the J-P production function).

In a reconciliation between the choice of LRP or a smooth production function, Berck and Helfand (1990) argued that the heterogeneity in the resource availability distribution over larger areas limit plant growth due to deficiency of one or more inputs. The heterogeneity of input availability through different type of inputs and quantity by which they are deficient, restricts plant growth differently, which leads to a smooth production function on the aggregate. Hennessy (2009) developed a micro-founded approach to yield distribution based on the distribution of resource availability following the arguments of Berck and Helfand (1990), and showed that LRP is capable of supporting both positive and negative yield skewness. Therefore, in this chapter a smooth (continuous and differentiable) stochastic production function is used to model the crop yield distribution.

The most popular choice of yield distribution has been the beta density (Day, 1965; Nelson and Preckel, 1989; Babcock and Hennessy, 1996; Hennessy, 2011; Zhu et al., 2011; Du et al., 2012) due to its flexibility to incorporate positive and negative skewness and its nature of double bounded domain (which is not unlikely about the crop yields that have a minimum and a maximum limit). Almost all studies that used a beta distribution have provided evidence in favor of a negatively skewed yield distribution for corn. Other distributions that have gained popularity among researchers in modeling the yield distribution include normal distribution (Fuller, 1965; Just and Weninger, 1999)[1], gamma distribution (Gallagher, 1986; Babcock and Blackmer, 1992)[2] and Johnson distribution (Ramirez et al., 2003). Few researchers have fit some popular distribution on transformed variables (like hyperbolic tangents, inverse hyperbolic sine) (Taylor, 1984; Moss and Shonkwiler, 1993; Ramirez, 1997). Non-parametric distribution have been used by Ker and Coble (2003). Few studies have compared the performance of different distributions to model yield (Norwood et al., 2004; Sherrick et al., 2004, 2014). While Norwood et al. (2004) argued in favor of semi-parametric distribution using an out of sample likelihood criteria, Sherrick et al. (2004, 2014) found beta and Weibull distributions to be the most appropriate. Based on the above discussion, beta density is assumed to be a reasonable description of the crop yield uncertainty and the same is adopted for modeling the yield density in this chapter

Early work of Thompson (1986, 1988) in agronomy estimated the effect of weather deviations from normal (long term average) on corn yield across five states. They found evidence of positive correlation of corn yields with normal pre-season (September through June) precipitation and above normal precipitation in July-August. They also reported yields to be positively correlated

[1] Fuller (1965) found normal distribution to be a good fit when carryover effects of previous nitrogen was removed.

[2] Babcock and Blackmer (1992) modeled the soil nitrate density conditional on which yield were modeled

with below normal temperatures in July-August respectively. Carlson (1990) found negative correlation of heat stress and positive correlation of plant available soil moisture with corn yields in central Iowa.

Agricultural economists have modeled the effect of weather on yield distribution (Kaylen and Koroma, 1991; Park and Sinclair, 1993). Kaylen and Koroma (1991) found evidence of a positive correlation of yield between July precipitation and negative correlation with August temperature. Their results also reported diminishing returns to weather variables. Park and Sinclair (1993) modeled the effect of weather variables on the moments of the yield distribution. Mean and variance of the yield distribution were negatively correlated to mean temperature. Moreover, increase in mean temperature reduced the negative skewness (made it less negative) of the yield distribution towards zero. Precipitation was positively correlated to the mean of the yield with no significant effect on the yield variation, and it increased the yield skewness (making the yield skew more negative). They also reported that warmer temperature along with low precipitation reduce yield skewness.

Chen et al. (2004) used the J-P production function and reported similar results of positive correlation between yield and precipitation, and negative correlation between yield and temperature. The effect on yield variability are reversed. In a recent conceptual model developed by Hennessy (2009), it was stated that a beneficial change in the distribution of weather variable indicated by an increase in the skewness of the weather distribution, causes a relatively greater increase in the skewness of the yield distribution (makes yield skewness more negative), if yields are characterized by diminishing marginal product of weather. Moreover, Hennessy (2011) showed that extreme weather events have less effect on the mean yield and the yield variability as their occurrence are closer to the harvest season.

Schlenker and Roberts (2006, 2009) found that the effect of temperature on the corn yield is non-linear and affected by the distribution of temperature around the average daily temperature (in growing degree days). Roberts et al. (2012) reviewed the role of agronomic variables in the crop growth process and identified growing degree days (GDD), heat degree days (HDD) and precipitation or vapor pressure deficit (VPD) to be the most critical agronomic weather variables in the plant growth process. Results from ordinary least squares (OLS) regression estimates showed evidence of negative correlation of yield variable with all weather variables except precipitation, to which it is positively correlated. They also report that increased precipitation mitigates the effect of HDD. Du et al. (2015) used the generalized method of moments (GMM) based on the flexible moments approach of Antle (1983), to model the moments of yield distribution conditional on the geographic location (latitude and longitude) and the climatic factors. They found similar evidence of increased mean yields with increased precipitation and decreased mean yields with increased overheated days. Overheated days make crop yield skew positive or less negatively skewed, and growing season precipitation reduce yield skewness (makes skewness more negative). In the present study, to account for the weather variables, growing season precipitation and growing season temperatures are used (as these two variables find significant mention in almost all studies discussed above). Pre-season precipitation is used as an indicator of available soil moisture during planting as described by Thompson (1986, 1988).

Most of the studies that use time series yield data have used some trend specification in their modeling to account for technological change over time. While Thompson (1986, 1988); Carlson (1990); Ramirez (1997); Chen et al. (2004); Roberts et al. (2012) used deterministic linear time trend[3], Kaylen and Koroma (1991); Moss and Shonkwiler (1993) have used a stochastic trend. Many studies have assumed a quadratic time trend (Schlenker and Roberts, 2009; Boyer et al., 2015). Just and Weninger (1999) emphasized that misspecification of trend could bias the results and advocated use of polynomial trend to choose the best fitting trend polynomial. Following Just and Weninger (1999) several studies have adopted polynomial trend that include Ramirez et al. (2003); Sherrick et al. (2004); Harri et al. (2009); Claassen and Just (2011).[4] A polynomial time trend is adopted in this study following Just and Weninger (1999).

[3]Thompson (1986, 1988) divided the time period in three sub-time periods and used a different linear trend for each sub-period.

[4]Claassen and Just (2011) uses a hybrid of polynomial trend method outlined in Just and Weninger (1999) and Atwood et al. (2003) to detrend the yield data

## 3.3 Model

Modeling of the yield distribution using a linear regression model under the assumption of additive error term imposes arbitrary structure on the error term.[5] Moreover, an assumption of normal distribution for the additive error term restricts the skewness of the yields to be zero. Using an alternative modeling technique, generalized linear model (GLM) to model the yield distribution allows the dependent variable to follow a non-normal density and do not contain an additive error term (McCullagh and Nelder, 1989; Lindsey, 1997; Smithson and Merkle, 2013). Unlike the linear regression model, GLM model the moments of the distribution of the dependent variable (following a non-normal density), conditional on the predictor variables. Using a distribution for the dependent variable that adequately captures the data generating process of the dependent variable, GLM estimate the moments of the yield distribution beyond the mean, which provides more flexibility to the GLM over the linear regression model with additive error. The deviation of the dependent variable from the estimated mean of the distribution of the dependent variable is the deviance term, which is captured in the estimation of the variance of the distribution for the dependent variable.

GLM is primarily described by its three features: (i) density of the dependent variable (ii) linear combination of the predictor variables (iii) link function that connect the moments of the distribution to the linear combination of predictor variables. The most common linear regression model with normally distributed errors, can be interpreted as a GLM with (i) a normally distributed dependent variable (ii) $\mathbf{X}\beta$ is a linear combination of predictor variables with weights of parameter vector $\beta$ as the coefficient and (iii) Identity link function, which connects the expectation of the dependent variable of the linear predictor. For a dependent variable $Z$, the standard multiple linear regression can be written as (Smithson and Merkle, 2013)

$$Z|\mathbf{X}, \beta \sim N(\mu, \sigma^2)$$

$$\mu = \mathbf{X}\beta$$

[5] In a linear regression $y = \mu + \epsilon$, for yields to be non-negative, $y \geq 0 \implies \epsilon \geq -\mu$, is a unreasonable restriction on the errors.

This is an example of a Gaussian GLM. As already mentioned earlier, it is important to note that unlike regression models, GLM do not have an error term in the linear predictor. More generally for any distribution $f(.)$,

$$Y|\mathbf{X}, \theta \sim f(\mathbf{\Theta})$$

and

$$\mathbf{\Theta} = \mathbf{G}(\mathbf{X}\theta)$$

where, $\mathbf{\Theta}$ is the vector of the moments of the distribution $f$ and $\mathbf{G}(.)$ is the vector of the link functions that connect the linear predictor to the moments of the $f$ distribution. Although, the example of a Gaussian GLM used an identity link function, the choice of different distribution require different link function. Different canonical link functions are used over the linear predictor, $\mathbf{X}\theta$ for different distributions, so that the values of the distribution moments lie in an admissible range. A valid link function must be monotonic and differentiable in linear predictor, and the derivative of a link function predicts the change in moments of the distribution following a change in the predictor variable (McCullagh and Nelder, 1989).

GLMs are widely used with exponential family of distributions including Poisson, Gamma, Binomial, Log normal, Exponential, Inverse Gaussian with one or more link functions to connect the first and the higher moments of the distribution to the linear predictor variables. Until recently, the beta distribution was not a popular choice of distribution in GLMs. The beta distribution is a generalized distribution from the exponential family of distribution and is parent distribution to several other distributions (i.e. under appropriate assumption it converges to several well known distributions) (McDonald and Xu, 1995). While Paolino (2001) and Ferrari and Cribari-Neto (2004) have popularized the use of conditional beta distributions, it has been used in agricultural economics literature earlier than that in Nelson and Preckel (1989) and Babcock and Hennessy (1996). Ferrari and Cribari-Neto (2004) has rather reparameterized it in terms of mean and variance of the distribution, popularizing it as beta regression to model rates and proportions. Smithson and Verkuilen (2006) extended it to model double bounded variables beyond the $(0, 1)$ interval, so that transformed variable using the lower and upper bound is contained in $(0, 1)$ interval. Ferrari and Cribari-Neto (2004) has contributed to the development of the theory of conditional beta distributions as beta regressions by providing diagnostics and asymptotic large sample properties to conduct inference.

A standard beta density for a variable $y$ (where $0 < y < 1$) is described by a location and a scale parameter, $\omega$ and $\nu$. The variable $y$ following a beta density with location and scale parameters $\omega$ and $\nu$ is denoted as $y \sim Be(\omega, \nu)$. The pdf of the standard beta density is

$$f(y, \omega, \nu) = \frac{\Gamma(\omega + \nu)}{\Gamma(\omega)\Gamma(\nu)} \cdot y^{\omega-1} \cdot (1-y)^{\nu-1}.$$

The first two moments of the beta density are,

$$E[y] \equiv \mu = \frac{\omega}{\omega + \nu}, \qquad V(y) \equiv \sigma^2 = \frac{\omega\nu}{(\omega+\nu)^2(\omega+\nu+1)}$$

The parameters of interest in a GLM model are the moments of the distribution of the variable $y$. The re-parameterization of the location and scale parameters of the beta density to the mean and precision parameters, lends it a natural interpretation in the GLM context. With the re-parameterization, the effect of a change in the predictor variable on the mean and variance of the yield distribution is relatively more intuitive than their effect on the location and scale parameter, $\omega$ and $\nu$ respectively. Re-parameterization of the density function assuming that the lower and upper bounds of the variable are known (Smithson and Merkle, 2013), gives

$$\varphi = \omega + \nu, \qquad \omega = \mu\varphi, \qquad \nu = (1-\mu)\varphi, \qquad V(y) \equiv \sigma^2 = \frac{\mu(1-\mu)}{(\varphi+1)}$$

Therefore, $y \sim Be(\omega, \nu)$ can be rewritten as $y \sim Be(\mu\varphi, (1-\mu)\varphi)$. Parameter $\varphi$ is the precision parameter, which is inversely proportional to variance (as can be seen in the term $V(y)$). As it is evident from the variance function, $\sigma^2$ tends to zero as $\mu$ tends to zero or one. The beta GLM has two sub-models, one each to model the mean, $\mu$ and precision $\varphi$ respectively. Let $\boldsymbol{x_i}$ and $\boldsymbol{w_i}$ be the predictor variables for the mean and precision sub-model for the $i^{th}$ observation. The mean and the precision sub model can be respectively written as

$$g(\mu_i) = \mathbf{x_i}\beta, \qquad h(\varphi_i) = \mathbf{w_i}\delta$$

$\beta$ and $\delta$ are parameters associated with the linear predictor, whereas $g(.)$ and $h(.)$ are the link function for the mean and precision sub-model respectively. The canonical link function for the mean is logit because it provides a value for the mean of the beta density between $(0, 1)$, which is admissible under the standard beta density. Similarly, the canonical link function for the precision is log because the precision parameter should be non-negative. The canonical logit link function for the mean and log function for precision parameter can be simplified and written as

$$\mu = \frac{exp(\mathbf{X}\beta)}{1 + exp(\mathbf{X}\beta)}, \qquad \varphi = exp(\mathbf{Z}\delta)$$

The parameters of the beta GLM are estimated via the maximum likelihood estimation. The log-likelihood of a single observation $i$ is given by (Smithson and Merkle, 2013)

$$L(y_i, \mu_i, \varphi_i) = \mathbf{\Psi}(\varphi_i) - \mathbf{\Psi}(\mu_i\varphi_i) - \mathbf{\Psi}(\varphi_i - \mu_i\varphi_i) + \mu_i\varphi_i log(y_i) + (\varphi_i - \mu_i\varphi_i) log(1 - y_i) - log(y_i) - log(1 - y_i)$$

where $\mathbf{\Psi}$ is the log-gamma function. The log likelihood function is given by the product of log-likelihood across observations as

$$\mathbf{L}(\mathbf{y}, \beta, \delta) = \prod_{i=1}^{n} L(y_i, \mu_i, \varphi_i)$$

where $n$ is the number of observation and $\mathbf{y}$ is the vector of yield realizations $(y_i)$. $\beta$ and $\delta$ are the vector of parameters for the mean and the precision sub-model.

## 3.4 Data

The yield data used is obtained from the annual corn yield and nitrogen experiments conducted at the research farms of Iowa State University.[6] The weather data is accessed from the Iowa Environment Mesonet.

Many studies have used experimental nitrogen trial data to model the yield density, which are managed by agronomists. One of the primary use of this data is to model the agronomic relationship between yield and inputs or to assess the effect of different farming practices on the yield. The studies are controlled experiment, which tries to isolate the effect an external factor (central to the study) on the yield. The nitrogen experiments have served as the basis of nitrogen recommendations made by agronomists to farmers. The yield data collected from the

[6] The annual study is led by Dr. John Sawyer at the Department of Agronomy, Iowa State University. I am thankful to Dr. Sawyer for providing the data.

nitrogen experiments are the realized yield values from the stochastic production function of the agricultural output with nitrogen and weather as the input. The use of experimental data to model the yields is objective in the sense that it is not influenced by any behavioral feedback of a farmer. It establishes the scientific relationship between the inputs and the output which is independent of any subjective assessment or endogenous choice of inputs.

The yield data used for this study is the experimental nitrogen trial data from four different research farms: Ames, Sutherland, Kanahwa and Nashua, managed by ISU.[7] Each farm is located in a different region of Iowa, which has varied soil characteristics and experience different weather. Based on the classification of twelve Major Soil Regions in Iowa[8], Ames farm in central Iowa and Kanahwa farm in northern Iowa are associated with a principal soil region of Clarion-Nicollet-Webster type soil. Sutherland in north western Iowa has a primary soil region of Galva-Primghar-Moody soil and Nashua in north eastern Iowa has predominantly the Kenyon-Clyde-Floyd soil type. Different soil types significantly vary in their productivity and soil moisture. They also differ in their ability to carry over nutrients from previous crops. Locationally different farms experience different weather (which may or may not be significantly different) that affects the yield productivity of the field. Both the soil type and field location are important in this chapter, as it allows to adequately control for these variables while comparing the estimates of the objective model to the subjective estimates in chapter 4. Conversation with the agronomists associated with this study provided additional information about the farm management practices observed in these experiments. The nitrogen experiments are controlled for different crop rotations, tillage practices, timing of nitrogen application so that the variation in the corn yields across farms are due to physical individual characteristics of the farm, weather uncertainty, unobserved technological changes and different levels of nitrogen application. Separate yield data series for Soybean-Corn (SC) and Corn-Corn (CC) crop rotation are obtained. Corn hybrid seeds are used on all farms. Average planting dates are from middle of April to end of May. Best Management Practices are implemented in each

[7]Summary report of this data for few years are available at the website of ISU Research and Demonstration farms.

[8]Detailed description of each soil type can be found on the website of ISU Research and Demonstration farms.

trial on each farm.

Table 3.1 Yield data series description

| Farm | Rotation | Nitrogen treatment (lbs./acre) | Years |
|---|---|---|---|
| Ames | SC | 0, 60, 120, 180, 240 | 2000-2013 |
| Ames | CC | 0, 60, 120, 180, 240 | 1999-2013 |
| Sutherland | SC | 0, 40, 80, 120, 160, 200, 240 | 2001-2013 |
| Sutherland | CC | 0, 40, 80, 120, 160, 200, 240 | 2000-2013 |
| Kanawha | SC | 0, 40, 80, 120, 160, 200, 240 | 2005-2013 |
| Kanawha | CC | 0, 40, 80, 120, 160, 200, 240 | 2006-2013 |
| Nashua | SC | 0, 40, 80, 120, 160, 200, 240 | 2005-2013 |
| Nashua | CC | 0, 40, 80, 120, 160, 200, 240 | 2005-2013 |

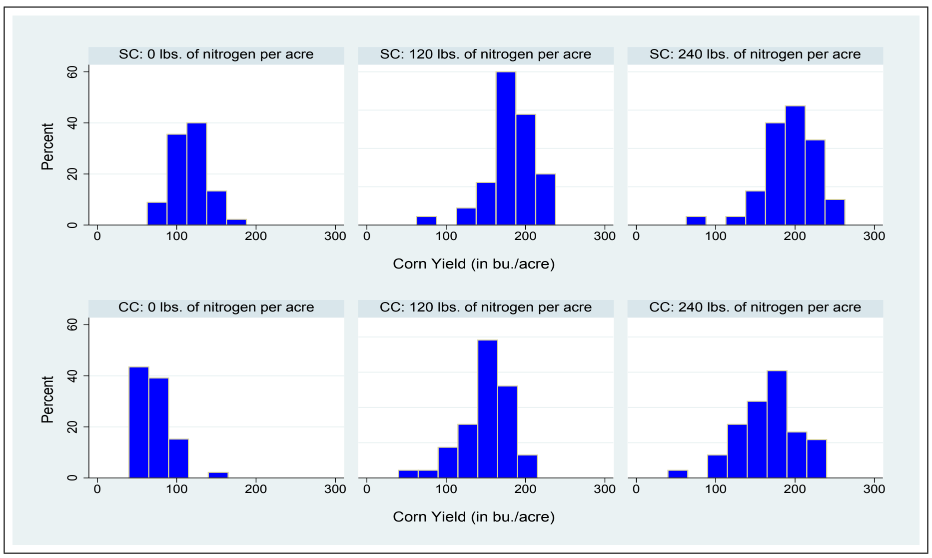



Figure 3.1 Histogram of corn yields conditional on three different level of nitrogen treatment: 0, 120 and 240 lbs./acre of nitrogen. SC and CC denotes soybean-corn and corn-corn rotation respectively

Table 3.1 describes the yield data series used from the nitrogen experiments. The level of nitrogen application treatment is from 0 to 240 lbs./acre in increments of 60 or 40 lbs./acre. For the Ames farm, yield data runs from the year 1999-2013 (2000 onward for SC). The yield data for Sutherland farm are available from the year 2000-2013 (2001 onward for SC). For both Kanwha and Nahsua farms, shorter yield time series from 2005-2013 (2006 onward for Kanawha CC) are available. Figure 3.1 plots the histogram of corn yield across three nitrogen treatment (0, 120 and 240 lbs./acre) and, SC and CC rotation. Figure 3.2 plots the corn yield over the years for Ames farm across rotation and the same nitrogen treatment.

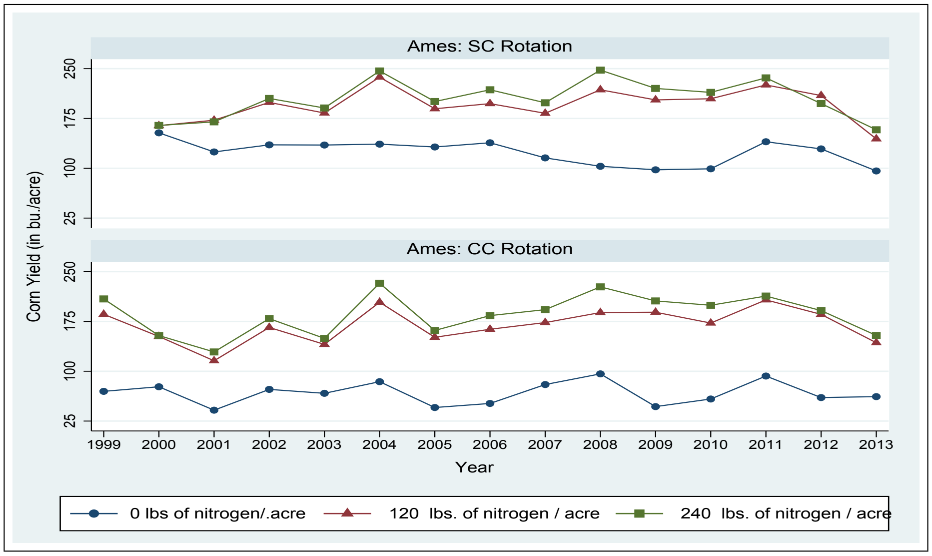



Figure 3.2 Time series of corn yields for Ames farm for three different level of nitrogen treatment: 0, 120 and 240 lbs./acre of nitrogen. SC and CC denotes soybean-corn and corn-corn rotation respectively

Weather data is obtained from the Climodat repository of the Iowa Environment Mesonet. Climodat is the climate data sourced from the National Weather Service Cooperative Observer Program, which is a nation-wide network of people making daily weather observations. Weather variables include daily precipitation (in inches), growing degree days (GDD), daily maximum and minimum temperatures (in degrees Fahrenheit). The ISU farms are mapped to the nearest available Climodat location to obtain the weather variables for the respective field.[9] Daily precipitation from May to August is used to create the variable for cumulative growing season precipitation. Similarly, daily precipitation from September to April is used to create cumulative pre-season precipitation. Using GDD and daily maximum and minimum temperature, heat degree days are calculated for a day, and cumulating the GDD and heat

[9] I am thankful to Daryl Herzmann, Assistant Scientist, Climate Science Program, Iowa State University for providing this information

degree days over the entire growing season from May to August, total degree days (TDD) is created. The summary statistic for weather variables are provided in table 3.2 below[10].

Table 3.2 Summary statistics of weather variables

| Variable | N | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Growing degree days | 47 | 2314.8 | 187.4 | 1838.0 | 2686.5 |
| Total degree days | 47 | 2371.9 | 213.0 | 1849.0 | 2843.5 |
| Growing season precipitation | 47 | 17.5 | 5.4 | 7.4 | 32.8 |
| Pre-season precipitation | 47 | 16.1 | 3.8 | 5.6 | 25.8 |

The linear predictor for both the mean and precision sub-model is assumed to be a flexible function of polynomials of nitrogen treatment, growing season precipitation, total degree days and pre-season precipitation and their interaction terms. Chebyshev polynomial bases functions are used for all continuous variables.[11] Separate fixed effects for site and rotation are introduced through site and rotation specific dummy variables. The agronomists involved in the nitrogen experiments mentioned that they adopt new technologies and better management practices over the years in the experiments. This leaves room for the effect of technical change on the yield distributions. Hence, following Just and Weninger (1999), a polynomial trend was included to account for technological changes.

## 3.5 Results

Conditional yield distributions are modeled using beta regression discussed in section 3.3. The parameter estimates of the beta regression model are reported in table 3.6 in the appendix in section 3.7. A flexible model of yield distribution is estimated. The model is flexible in the sense that it does not put unwarranted structure on the moments of the distribution. The flexibility in the model was allowed by using polynomials of nitrogen, weather and trend variables, and their interactions in the linear predictor of the GLM, which captured the effect

[10] It includes weather data for 15 growing season for Ames (1999-2013), 14 growing season for Sutherland (2000-2013), 9 growing season for Kanawha and Nashua (2005-2013)

[11] Use of Chebyshev polynomial is common while working with higher degree polynomial function as they reduce collinearity among polynomials of higher degree by forming an orthogonal basis, and hence stable parameter estimates particularly near the endpoints (Miranda and Fackler, 2004; Schlenker and Roberts, 2006, 2009).

Table 3.3 Summary of weather variables for Ames farm

| Variable | Cum. Growing Season Precp. | Total Degree Days | Cum. Preseason Precp. |
|---|---|---|---|
| N | 15 | 15 | 15 |
| $5^{th}$ percentile | 9.8 | 2151.5 | 7.6 |
| $10^{th}$ percentile | 13.0 | 2194.0 | 11.2 |
| $25^{th}$ percentile | 14.3 | 2381.5 | 14.2 |
| $50^{th}$ percentile | 18.3 | 2534.5 | 16.6 |
| $75^{th}$ percentile | 22.4 | 2616.5 | 18.7 |
| $90^{th}$ percentile | 30.5 | 2734.0 | 20.5 |
| $95^{th}$ percentile | 32.8 | 2686.5 | 25.8 |

of nitrogen, weather and time trend on moments of the yield distribution. Moreover, the estimation of a precision sub-model along with the mean sub-model, added to the flexibility of the model.

The three weather variables used in modeling the yield distribution is the growing season precipitation, TDD to account for temperature, and pre-season precipitation for soil moisture. Corn is a water intensive crop, so growing season precipitation is used in all the three models presented in table 3.6 in the appendix in section 3.7. Model 1 and model 2 used either pre-season precipitation or TDD respectively, while model 3 used both of them along with the growing season precipitation in estimating the yield distribution.

Variable selection in each of the three model is based on backward elimination using Wald test with heteroscedasticity consistent robust (sandwich estimates) standard errors. The Bayesian information criteria for model 3 is the smallest, which is used as a selection criteria to choose model 3 for illustration of results of the effect of nitrogen, weather variables and their interaction on the yield distribution. In model 3, which included all the three weather variables, the functional form chosen is cubic in nitrogen, as Wald test could not reject the null hypothesis that nitrogen terms beyond cubic are not significantly different from zero.

The first three moments of the yield distribution: mean yield, yield variability (standard deviation) and yield skewness for the Ames farm for the year 2014 are presented as a function of nitrogen and weather using estimates from model 3.

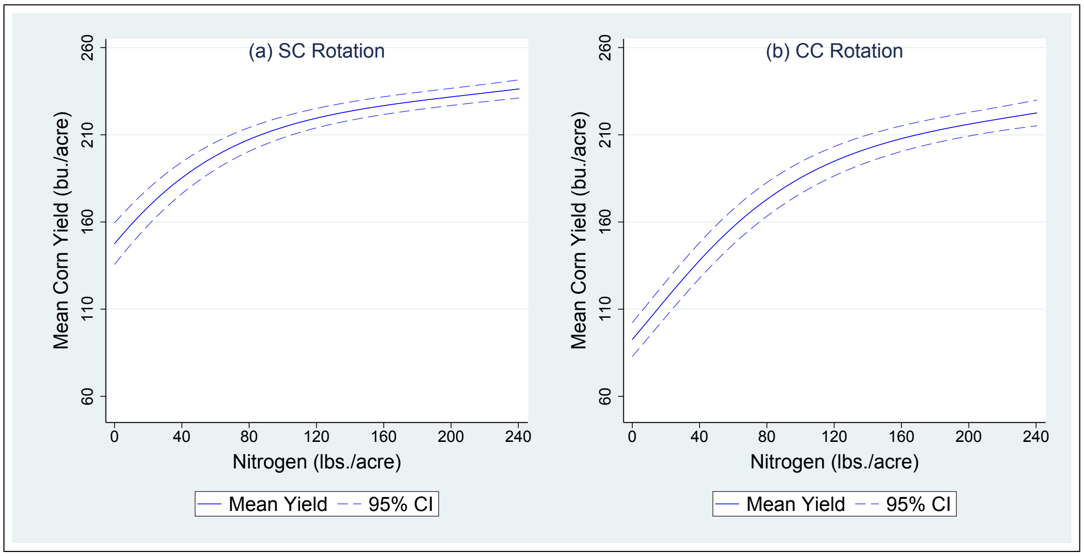



Figure 3.3 Mean Yield conditional on nitrogen and weather for the Ames farm: (*50%-50%-50%*) denotes a weather scenario with the $50^{th}$ percentile of the growing season rainfall, the TDD and the pre-season precipitation respectively.

The next few subsections are devoted to the results of this chapter, which summarizes the effect of nitrogen and weather on the yield distribution, and their implication for nitrogen management. For all graphs, the horizontal axis measures nitrogen application (with 0 to 240 lbs./acre of nitrogen). Table 3.3 summarizes the weather variable used for the Ames site. A *(P%-Q%-R%)* format has been adopted to depict a particular weather scenario. A weather scenario *(P%-Q%-R%)* corresponds to $P^{th}$ percentile of the growing season precipitation, $Q^{th}$ percentile of the TDD and $R^{th}$ percentile of the pre-season precipitation. For example, a weather scenario with median weather variables denoted by *(50%-50%-50%)* indicates a weather condition (from table 3.3) in which 18.3 inches of growing season precipitation, 2534.5 degree days of growing season temperature (from May to August) and 16.6 inches of pre-season precipitation (from September to April) has occurred.

### 3.5.1 Mean Yield and Marginal productivity of nitrogen

#### Mean Yield

Figure 3.3 plots the weather and nitrogen conditional mean yield at all levels of nitrogen (from 0 to 240 lbs./acre) for the SC (panel (a)) and the CC rotation (panel (b)). The weather condition at which the mean yield is evaluated is the median value of all the weather variables. The mean yield is substantially higher for the SC rotation relative to the CC rotation at low levels of nitrogen, but the yield gap decrease as the amount of nitrogen application increase.

#### Expected Utility / Profit maximization

If farmers are believed to be maximizing utility or profit, the marginal product of nitrogen (MPN) is instrumental in the nutrient decision-making. Under utility maximization under uncertainty, a farmer's expected utility is given by

$$\max_N \; EU = \int U(\pi(N, W)) \cdot dF(W)$$

where, $EU$ is the expected utility and $U(.)$ is the utility function. $\pi(N, W)$ is the farm profit such that

$$\pi(N, W) = P \cdot f(N, W) - P_N \cdot N$$

$f(N, W)$ is the yield corresponding to $N$, nitrogen and $W$, weather. $P$ and $P_N$ are the corn and nitrogen prices respectively. $F(W)$ is the cumulative joint probability distribution of weather variables.

Similarly, for a profit maximizing farmer, expected profit is given by

$$\max_N \; \pi(N) = \int \pi(N, W) \cdot dF(W)$$

In expected utility framework, $U(\pi)$ is the farmer's utility over profits, such that $U_\pi > 0$. A farmer is said to be risk-neutral if $U_{\pi\pi} = 0$, and risk-averse if $U_{\pi\pi} < 0$. Under the above framework, optimal nitrogen $N^*$ is determined by

$$\int U_\pi \cdot \frac{\partial \pi(N^*, W)}{\partial N} \cdot dF(W) = 0 \implies P \cdot \int U_\pi \cdot f_N(N^*, W) \cdot dF(W) = P_N \cdot \int U_\pi \cdot dF(W)$$

Similarly, under expected profit maximization, the producer's optimal choice of nitrogen is $N^*$ such that

$$\int \frac{\partial \pi(N^*, W)}{\partial N} \cdot dF(W) = 0 \implies P \cdot \int f_N(N^*, W) \cdot dF(W) = P_N$$

For a risk neutral farmer, $U_\pi$ is a constant, and therefore both the expected utility and expected profit maximization leads to same level of optimal nitrogen given by

$$\int \frac{\partial f(N^*, W)}{\partial N} \cdot dF(W) = \frac{P_N}{P} \tag{3.1}$$

for any given corn and nitrogen price ratio.

Under the expected profit maximization or expected utility maximization under risk neutrality as in equation 3.1, four different pieces of information are needed to arrive at a farmer's optimal amount of nitrogen application, which includes the nitrogen and corn prices, and the weather conditioned MPN and the distribution of weather variables. From a farmer's perspective under uncertainty, the corn price, the weather conditioned MPN and the weather distribution (weather uncertainty) is their subjective expectations of each of this variable. The choice of optimal nitrogen is determined by several pieces of information brought together as in equation 3.1. In this chapter objective estimates of expected MPN are calculated using agronomy farm data, whereas the next chapter measures the subjective estimates of expected MPN (the left hand term of equation 3.1), which is an essential part of the nitrogen decision-making problem. However, expected utility maximization with risk averse individuals additionally require information about the risk preference (utility function) of a farmer.

### Mean marginal productivity of nitrogen

In this chapter, the MPN of the nitrogen is estimated from the experimental nitrogen trial data of the agronomy farms. The optimal nitrogen choice of an expected profit or a risk neutral expected utility maximizing farmer is derived for a hypothetical farmer for a feasible corn-nitrogen price ratio.

Figure 3.4 is the graph of MPN across SC and CC rotation, estimated at the median value of the weather variables. MPN is positive across both the rotations at all levels of nitrogen

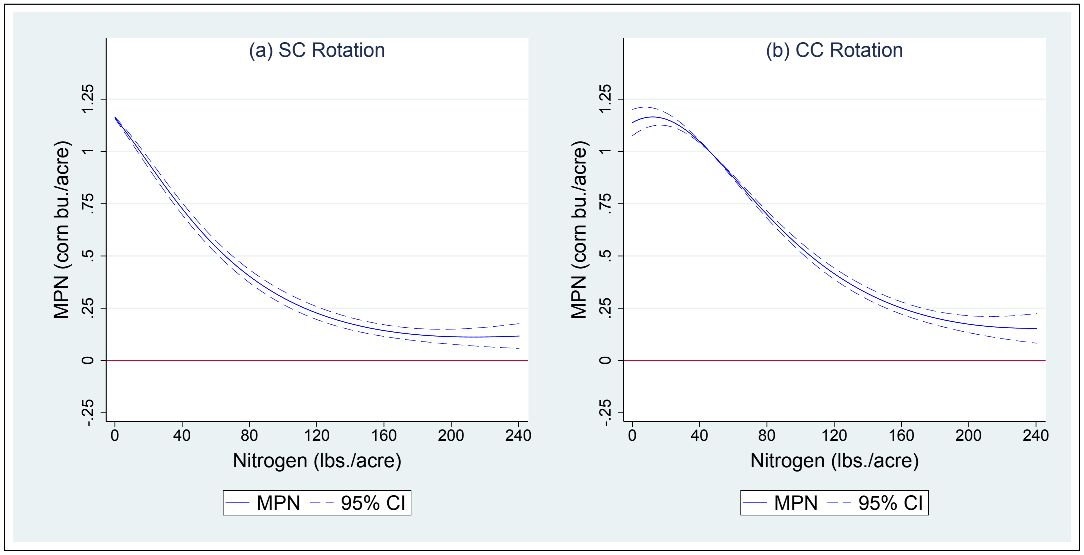



Figure 3.4 MPN conditional on nitrogen and weather for the Ames farm: (*50%-50%-50%*) denotes a weather scenario with $50^{th}$ percentile of growing season rainfall, growing season degree days and pre-season precipitation respectively.

applied. The downward sloping curve of MPN validates that the results are consistent with the positive and diminishing MPN (with the exception of CC rotation where MPN is non-decreasing at initial low levels of nitrogen application).

For an expected profit maximizing farmer or a risk neutral expected utility maximizing farmer, changes in weather condition may affect the MPN, and subsequently the choice of optimal nitrogen. Figures 3.5, 3.6 and 3.7 compares the MPN under different weather scenario and their implications for choice of optimal nitrogen for an expected utility maximizing risk neutral farmer.

For a hypothetical farmer, whose expected MPN is assumed to be the MPN corresponding to (50%-50%-50%) weather condition (as in Figure 3.4)[12], the optimal choice of nitrogen is determined by equating MPN to the price ratio of nitrogen and corn. If the nitrogen-corn price ratio is 0.10 ($0.50 : $5.00), the optimal nitrogen will be around 180 and 210 lbs./acre for SC and CC rotation respectively.

[12]This holds only under the assumption that at all levels of nitrogen, the expected MPN is same as the MPN under the median weather scenario

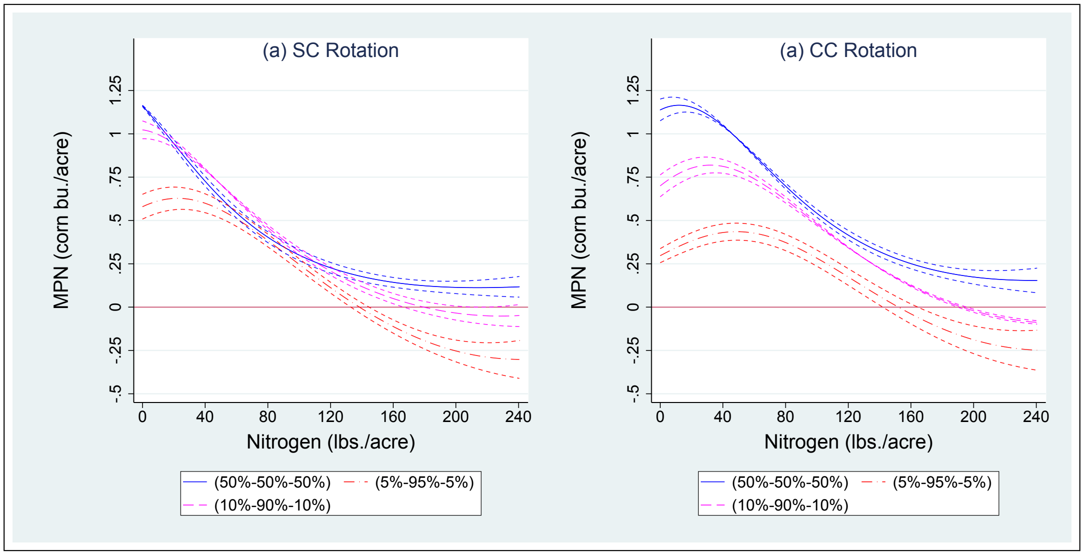



Figure 3.5 MPN conditional on nitrogen and weather for the Ames farm. The MPN correspond to (50%-50%-50%), (5%-95%-5%) and (10%-90%-10%) weather scenarios.

In figure 3.5, the MPN under median weather conditions are compared to extremely hot and dry growing weather conditions with low pre-season precipitation denoted by (5%-95%-5%) and (10%-90%-10%). In the above two scenarios, each of this denotes low pre-season precipitation, which is indicative of low soil moisture before planting. Moreover, low soil moisture is accompanied with extreme heat and high growing season water stress. The percentiles of the weather conditions are described in table 3.3. First $5^{th}$ and $10^{th}$ percentiles denote 9.8 and 13 inches of growing season precipitation, followed by $95^{th}$ and $90^{th}$ percentiles of heat equivalent to 2686.5 and 2676.5 TDD respectively. The third value is $5^{th}$ and $10^{th}$ percentiles of pre-season precipitation, which is equivalent to 7.6 and 11.2 inches of rainfall. It can be seen in figure 3.5 that MPN is significantly low and can also become negative under extreme heat stress such as the $90^{th}$ percentile (at 2734 units) of TDD or higher. Corresponding to a nitrogen-corn price ratio of 0.10, the choice of optimal nitrogen is significantly lower if the expected growing season weather is relatively hot and dry, preceded by low pre-season precipitation. The optimal nitrogen corresponding to (10%-90%-10%) is 160 lbs./acre for SC rotation and 170 lbs./acre for CC rotation. The optimal nitrogen levels are even lower, at 130 lbs./acre and 140 lbs./acre for SC and CC rotation respectively for more adverse weather denoted by (5%-95%-5%).

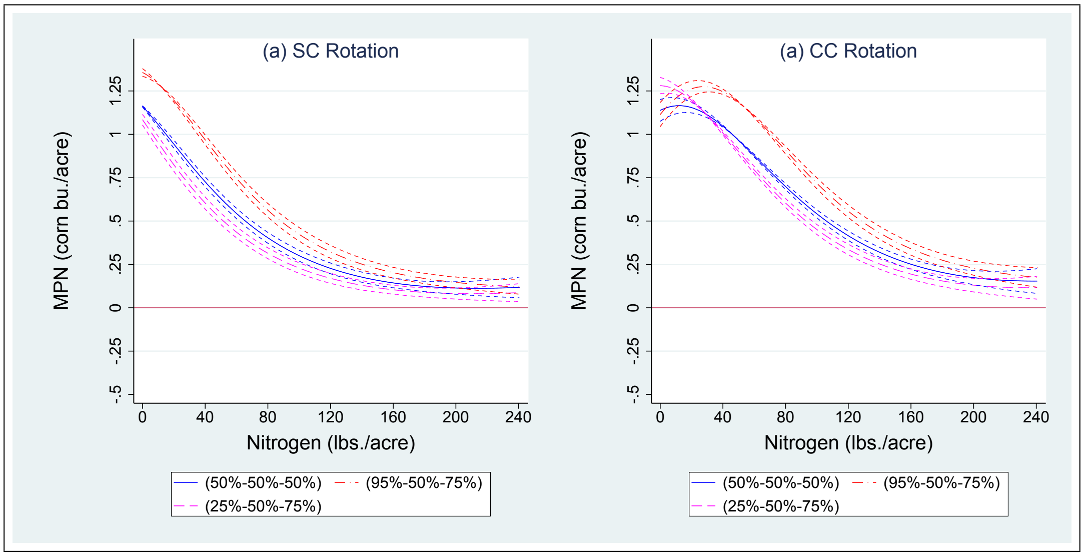



Figure 3.6 MPN conditional on nitrogen and weather for the Ames farm. The MPN correspond to (50%-50%-50%), (95%-50%-75%) and (25%-50%-75%) weather scenarios.

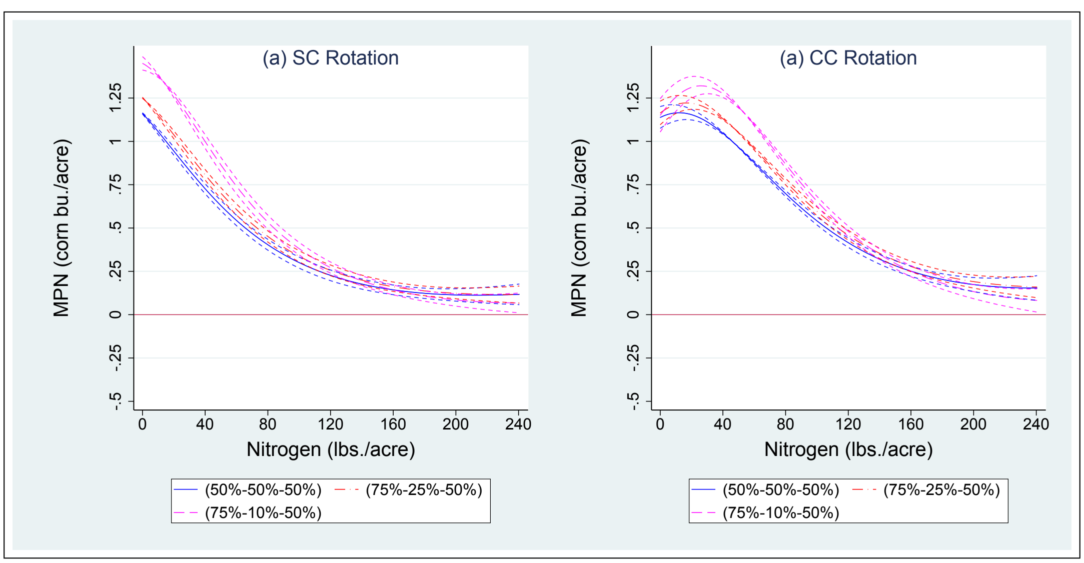



Figure 3.7 MPN conditional on nitrogen and weather for the Ames farm. The MPN correspond to (50%-50%-50%), (75%-25%-50%) and (75%-10%-50%) weather scenarios.

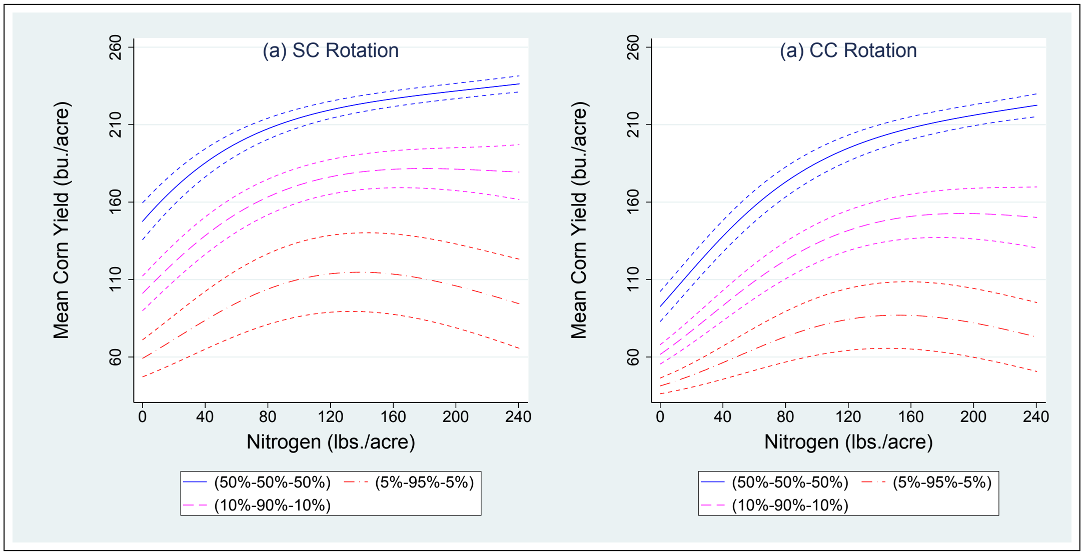



Figure 3.8 Mean Yield conditional on nitrogen and weather for the Ames farm. Mean Yield correspond to (50%-50%-50%), (5%-95%-5%) and (10%-90%-10%) weather scenarios.

In figures 3.6 and 3.7, MPN under several other weather scenarios are plotted. The weather scenarios comprise of situation where the weather conditions are more (less) favorable than the median scenario through moderate increase (decrease) in the pre-season or growing season rainfall or a moderate reduction in the growing season heat. The weather scenarios are denoted by (95%-50%-75%) and (25%-50%-75%) in figure 3.6, and by (75%-25%-50%) and (75%-10%-50%) in figure 3.7 respectively. Weather change results in significant changes in the MPN, but the change is insignificant for nitrogen higher than or equal to 100 lbs./acre (approximately). In a feasible range of nitrogen-corn price ratio (say less than 0.20), the weather change implies that the choice of optimal nitrogen remains unaffected as there is no significant change in the MPN in this range. Therefore, optimal nitrogen is responsive to extreme weather (very hot and dry season) only, and under all other weather change it remains unaffected.

Figure 3.8 compares the mean yield for (10%-90%-10%) and (5%-95%-5%) weather scenarios with the mean yield under median weather scenario. It can be seen that both weather scenarios (10%-90%-10%) and (5%-95%-5%) represent strongly stressful situation as the average yields are significantly depressed as compared to the median weather scenario. Besides the extreme dry and hot season, average yields are not affected significantly under other weather scenarios.

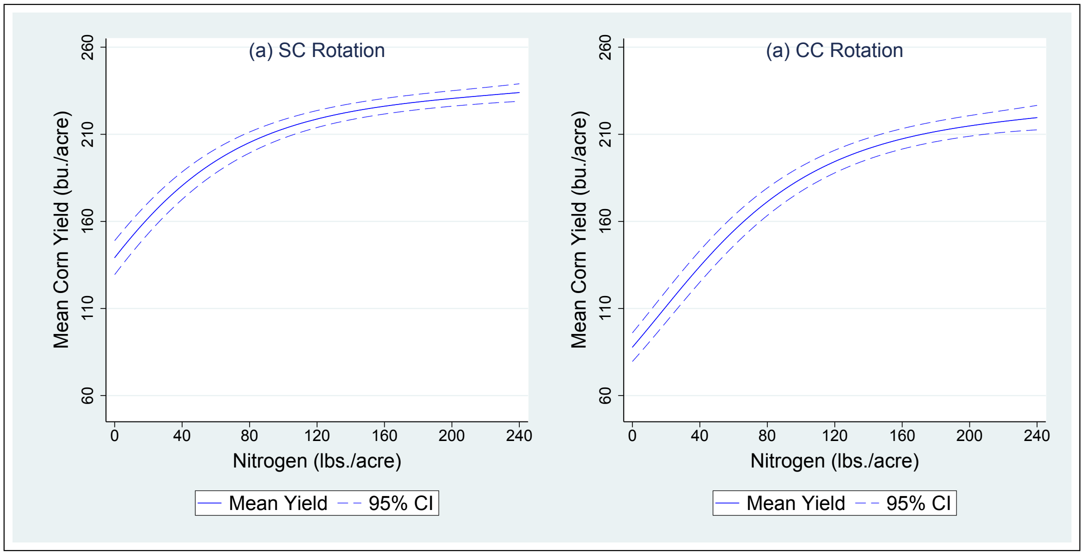



Figure 3.9 Avearge yield conditional on nitrogen for the Ames farm under weather uncertainty.

#### Mean Yield and MPN under weather uncertainty

The preceding discussion regarding the average yield and the MPN has been in context to a specific weather scenario, where it has been assumed that the expected MPN (over weather) at all levels of nitrogen is described by the MPN under a specific weather scenario. Under this assumption, the relevance of a particular weather condition in context to optimal nitrogen choice is discussed. However, more generally, weather uncertainty expectations of a farmer are based on rational expectations. Therefore, for a representative farmer the yield and the MPN is weighted over the distribution of the weather variables based on rational expectations.

In order to derive rational expectations of average yield and average MPN, a distribution over the weather variables is fitted. The vector of weather variables (which include pre-season precipitation, growing season precipitation and growing season TDD) is assumed to follow a joint log normal distribution. Over the domain of the Ames weather described in table 3.3, a joint trivariate log normal distribution is fitted. The probabilities of the estimated trivariate log normal weather distribution are used as weights to get the mean yield and mean MPN under the assumed weather uncertainty.

Figures 3.9 and 3.10 plots the average yield and average MPN respectively over SC rotation (panel a) and CC rotation (panel b). A trivariate joint log-normal distribution was fitted over

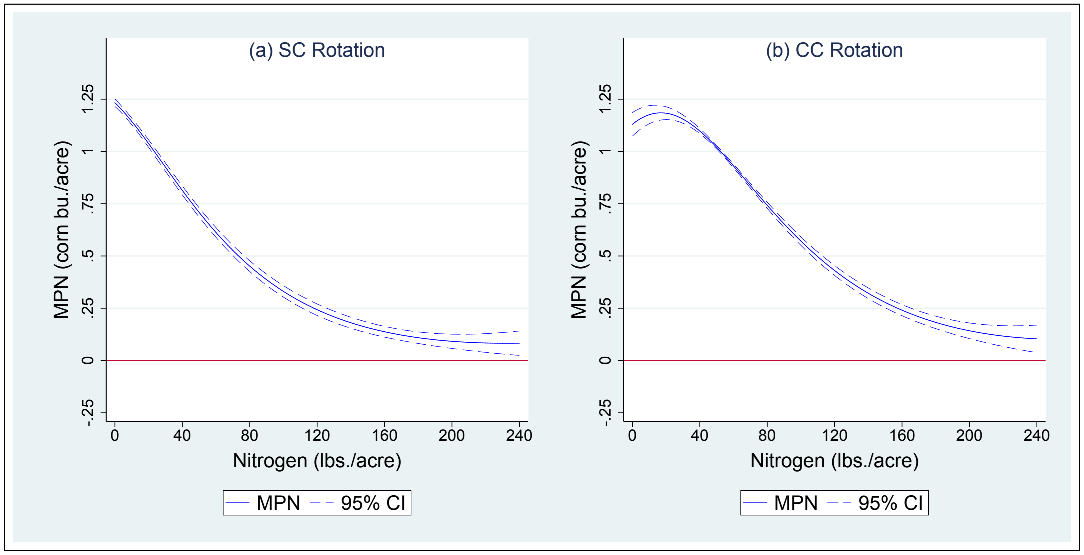



Figure 3.10 Average MPN conditional on nitrogen for the Ames farm under weather uncertainty.

the three weather variables, growing season precipitation, growing season heat and pre-season precipitation over the fifteen years of weather data for Ames farm. For a risk neutral expected utility or an expected profit maximizing farmer, figure 3.10 implies that the optimal nitrogen application for SC rotation is 180 and for CC rotation is 210 lbs./acre at 0.10 nitrogen-corn price ratio. This holds under the assumption that the farmer believes the weather uncertainty to be represented by the joint log normal distribution over growing season precipitation, growing season heat and pre-season precipitation. Moreover, it is also assumed that the MPN for the Ames farm adequately capture their beliefs about MPN.

### Convexity of MPN

Babcock and Blackmer (1992) have argued that convex MPN curve could be a potential reason for nitrogen over application by farmers. If the farmer's share similar beliefs about the shape of MPN, they might operate under the rule of thumb to apply more fertilizer in case it is needed due to greater optimal levels of nitrogen under uncertainty, because average gain of an additional unit of nitrogen is higher than the average loss.

MPN is convex in nitrogen, if the second derivative of MPN is greater than zero. Under weather uncertainty this implies that $\dfrac{d^3(\int f(N,W)\cdot dF(W))}{dN^3} > 0$. In all the figures for the

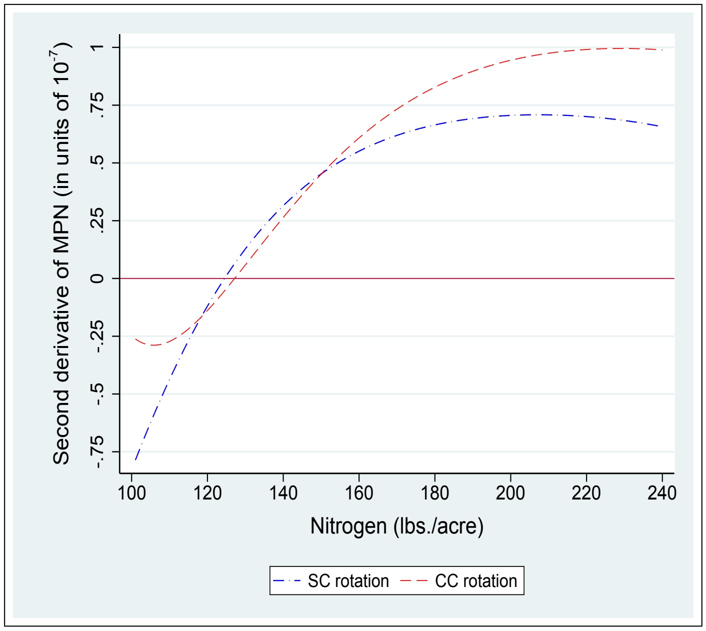



Figure 3.11 Curvature of MPN (second derivative of MPN) conditional on nitrogen for the Ames farm under weather uncertainty.

MPN (3.4, 3.5, 3.6, 3.7 and 3.10) plotted across nitrogen, it is observed that the MPN is convex in the operational range of a typical farmer (range of optimum nitrogen that corresponds to less than 0.20 of nitrogen-corn price ratio). Figure 3.11 plots the second derivative of the MPN with respect to nitrogen. In the range 100 to 240 lbs./ace of nitrogen application, the second derivative of MPN is positive, which indicates the possibility of convex MPN in this range (although this is not tested).

### Plateau function of yield nitrogen relationship

Evidence of positive and diminishing MPN is found in figures that illustrate the relationship between MPN and nitrogen. Moreover, constancy of MPN (which may be even zero (or close to zero) at higher levels of nitrogen application is suggestive of existence of plateau response of yield to nitrogen. Figures 3.5 to 3.7 indicate that the level of nitrogen beyond which MPN is zero depend on the weather variables and their interaction with nitrogen. The uncertainty of the weather variable could be the reason for stochastic plateau yield models to perform better than their non-stochastic counterpart in terms of model fit and performance (Tembo et al.,

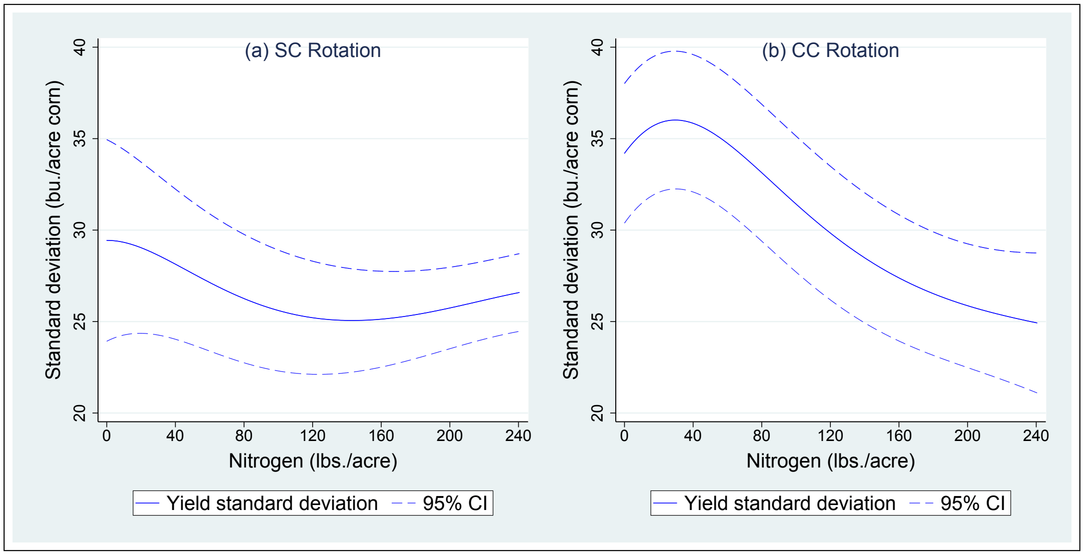



Figure 3.12 Standard deviation of yield conditional on nitrogen and weather for the Ames farm: (*50%-50%-50%*) denotes a weather scenario with the $50^{th}$ percentile of growing season rainfall, the TDD and the pre-season precipitation respectively.

2008; Tumusiime et al., 2011; Boyer et al., 2013).

### 3.5.2 Yield Variance

#### Yield variance under median weather conditions

Estimates of the moments of the yield distribution in table 3.6 in the appendix, section 3.7 include weather variables in the model. The yield variability is conditional on the weather and nitrogen variables, and therefore the yield variability is caused by unaccounted weather conditions or other factors like the physical characteristics of the field including the nutrient carry over from previous growing season, planting and pollination dates etc. Figure 3.12 plots the standard deviation of yields at the median weather scenario (i.e. at the $50^{th}$ percentile of the growing season precipitation, the TDD and the pre-season precipitation). It can be seen that the yield variability is not significantly different for the SC and CC rotation at the 95% confidence intervals.

**Yield variance under weather uncertainty**

As already mentioned earlier, the estimates of the yield standard deviation provided in the previous section are conditional on nitrogen and weather. Therefore, the yield variability cannot be characterized as a significant amount of weather uncertainty has been netted out. In order to get at the yield variance of nitrogen input, the yield variance is integrated over the probability weighted domain of weather variables to arrive at the yield variance conditional on nitrogen only, in which the aggregate weather uncertainty is present.

Let $y|N, W$ denote the nitrogen ($N$) and weather($W$) conditional yield variable. The first two moments of the yield distribution conditional on weather and nitrogen are given by:

$$Expectation: E[y|N,W] \quad Variance: V[y|N,W] = E[y^2|N,W] - (E[y|N,W])^2 \tag{3.2}$$

The first two moments of the yield distribution conditional on nitrogen only, under the assumption of an aggregate weather uncertainty is given by:

$$Expectation: E[y|N] \qquad Variance: V[y|N] = E[y^2|N] - (E[y|N])^2 \tag{3.3}$$

Let the weather uncertainty be denoted by a random variable that follows a cumulative distribution such that $W \sim F(W)$ over the weather domain $(\underline{W}, \overline{W})$. Subsequently, the relationship between the moments of the two models (conditional upon nitrogen only and conditional upon nitrogen and weather) can be derived.

$$E[y|N] = \int E[y|N,W] \cdot dF(W) \qquad E[y^2|N] = \int E[y^2|N,W] \cdot dF(W)$$

$$V[y|N] = \int E[y^2|N,W] \cdot dF(W) - \left( \int E[y|N,W] \cdot dF(W) \right)^2$$

Using the variance of both nitrogen and weather conditional yield,

$$E[y^2|N,W] = V[y|N,W] + (E[y|N,W])^2$$

$$\int E[y^2|N,W] \cdot dF(W) = \int V[y|N,W] \cdot dF(W) + \int (E[y|N,W])^2 \cdot dF(W)$$

Using this in the equation for nitrogen only conditional yield,

$$\begin{aligned} V[y|N] = &\int V[y|N,W] \cdot dF(W) + \int (E[y|N,W])^2 \cdot dF(W) \\ &- \left( \int E[y|N,W] \cdot dF(W) \right)^2 \end{aligned} \tag{3.4}$$

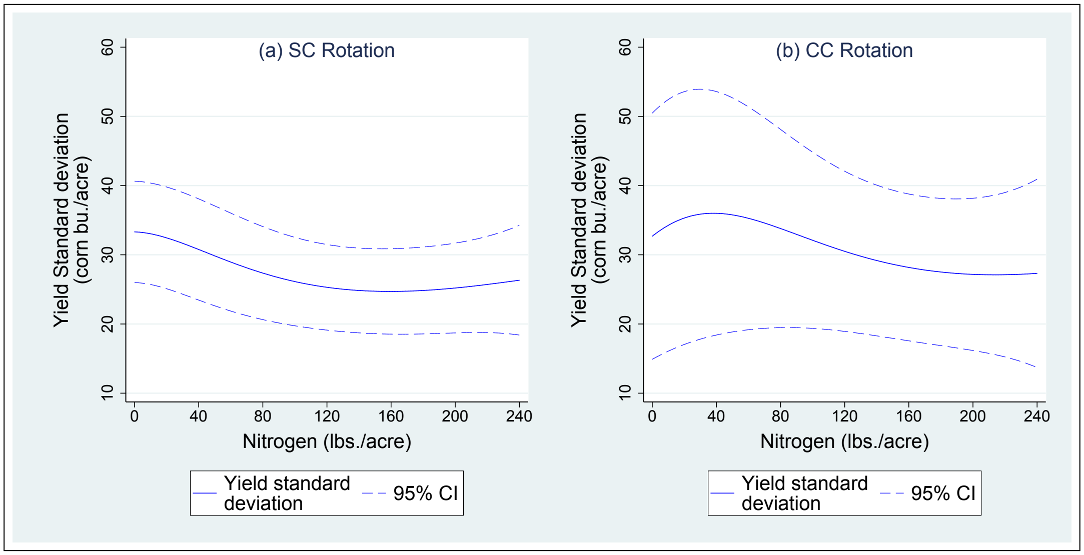



Figure 3.13 Yield standard deviation conditional on nitrogen for the Ames farm under weather uncertainty.

A model of yield distribution conditional on nitrogen, which does not factor out the effect of weather variables, implicitly assumes that the residual represents the yield variability due to weather. The estimated yield variance assumes an aggregate weather uncertainty that is captured by a mean weather scenario. Therefore, instead of estimating the yield variability at a mean weather scenario, the entire weather distribution is integrated back into the estimates of nitrogen and weather conditional yield mean and the variance to arrive at nitrogen conditional yield variance as described in equation 3.4. Previously estimated trivariate joint log normal weather distribution that was used for estimating average yield and MPN under weather uncertainty is used for estimating the weather specific variance.

Figure 3.13 plots the nitrogen conditional yield standard deviation as a function of nitrogen under the specified weather uncertainty. It can be seen that across both SC and CC rotation, yield standard deviation is fairly constant across amount of nitrogen applied. These results may be contradictory to earlier research (Just and Pope, 1979), which have shown that the yield variance increase as amount of nitrogen application increase. It can be seen in figure 3.13 that yield standard deviation is not necessarily increasing in nitrogen.

### Yield variance as risk increasing input

Researchers have considered variability of the yield distribution synonymous to yield risk and has concluded nitrogen to be a risk increasing input as they found nitrogen to increase the variance of the yield distribution (Just and Pope, 1979). Figure 3.13 provides evidence against this claim and shows that nitrogen is not necessarily a risk increasing input.

Rothschild and Stiglitz (1970) suggest that although variance is used as a traditional measure of risk by adopting a mean-variance approach, it is a spurious measure of riskiness of an uncertain outcome and could be misleading. They provide three alternative but equivalent intuitive characterization of riskiness of a random variable. One of three concepts of risk involved comparing two random variables, such that more weight in tails of the distribution represent a riskier random variable. A second concept defines a random variable to be less risky if it is preferred by every risk averter (those with concave utility functions). The third one describes the new random variable to be riskier, if the new random variable is generated by adding an uncorrelated zero mean random variable. An important condition for all the above definitions including the variance as a measure of risk is that for all of them, both the random variables compared have the same mean.

The prevalence of the mean-variance approach to measure riskiness of a random variable is due to its ability to provide a complete ordering, which the other three equivalent approaches do not, and are capable of providing only a partial ordering. Nonetheless, under adequate assumptions Rothschild and Stiglitz (1970) showed that the riskiness of a random variable can be correctly determined by the three alternative definitions of risk unlike the mean-variance approach, which can rank an outcome to be riskier (due to higher variance) in spite of the fact that it is less risky.

In context to nitrogen and yield risk, SriRamaratnam et al. (1987) reported evidence of lower yield variance up to certain level of nitrogen from farmer's nitrogen beliefs. This implies that most farmers believe nitrogen to be a risk reducing input. Moreover, Paulson and Babcock (2010) have found evidence from their survey, where most farmers have reported nitrogen to

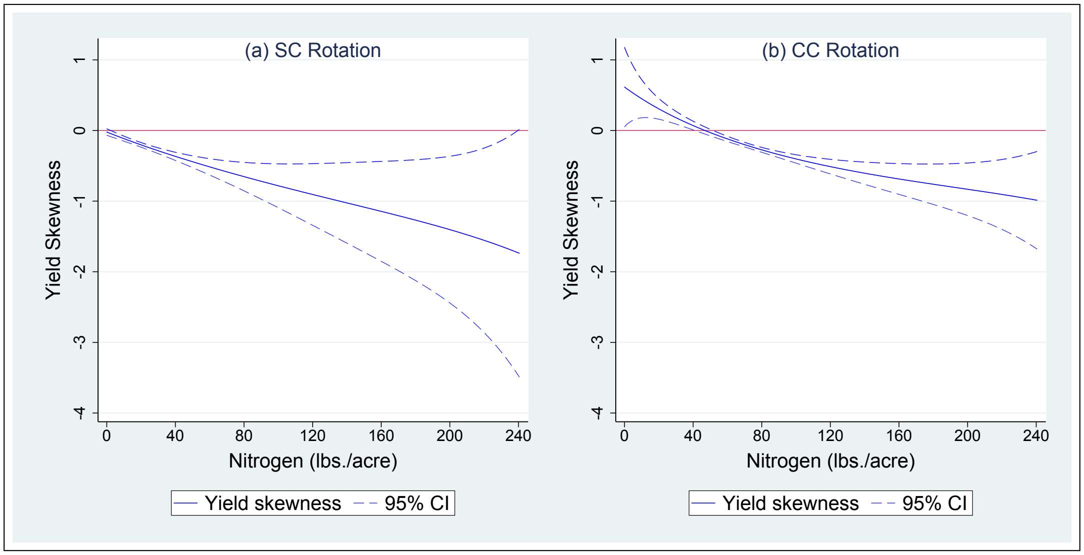



Figure 3.14 Skewness of yield conditional on nitrogen and weather for the Ames farm: (*50%-50%-50%*) denotes a weather scenario with $50^{th}$ percentile of growing season rainfall, growing season degree days and pre-season precipitation respectively.

increase the yield variance, but they believe nitrogen to reduce risk. This is an indication to the fact that farmers do not believe yield variance to be a measure of yield risk.

### 3.5.3 Yield Skewness

Yield skewness measures the asymmetry of the yield distribution. A measure of yield skewness weighs the cumulative probability for yield less than mean with the cumulative probability for yield greater than mean. If less than probability is greater (less) than greater than probability, the distribution is positively (negatively) skewed.

Previous research studies have found that the yield density is positively skewed at low levels of nitrogen application, and as nitrogen application increase, the yield skewness becomes symmetric and eventually changes sign to negative (Day, 1965; Nelson and Preckel, 1989; Babcock and Hennessy, 1996; Du et al., 2012). At even higher nitrogen application, yield skewness increases (becomes more negative).

Similar results of yield skewness are reported from the estimated model of yield density. In figure 3.14, yield skewness is plotted against nitrogen for a weather condition denoted by the median values of the weather distribution. It can be seen that for both SC and CC rotation, yield skewness increase (becomes more negative), as nitrogen application increase. Although, in the SC rotation, yield skewness is mostly zero or negative, CC rotation exhibit some positive yield skewness, which becomes negative around 60 lbs./acre of nitrogen application and continues to be negative.[13]

#### Yield skewness under weather uncertainty

Similar to the yield variance, yield skewness is also evaluated under weather uncertainty rather than at any specific weather scenario as in figure 3.14.

The skewness of the yield distribution conditional on nitrogen only are given by:

$$Sk[Y|N] = \frac{E[(Y - E[Y|N])^3|N]}{(V[Y|N])^{3/2}}$$

$$= \frac{1}{(V[Y|N])^{3/2}} \cdot \left(E[Y^3|N] - 3 \cdot E[Y^2|N] \cdot E[Y|N] + 2 \cdot E[Y|N]^3\right)$$

$$= \frac{1}{(V[Y|N])^{3/2}} \cdot \left(E[Y^3|N] - 3 \cdot (V[Y|N] + E[Y|N]^2) \cdot E[Y|N] + 2 \cdot E[Y|N]^3\right)$$

$$Sk[Y|N] = \frac{1}{(V[Y|N])^{3/2}} \cdot \left(\int E[Y^3|N,W] \cdot -3 \cdot V[Y|N] \cdot E[Y|N] - E[Y|N]^3\right) \tag{3.5}$$

Similarly, skewness of the yield distribution conditional on weather and nitrogen is:

$$Sk[Y|N,W] = \frac{E[(Y - E[Y|N,W])^3|N,W]}{(V[Y|N,W])^{3/2}}$$

$$= \frac{1}{V[Y|N,W]^{3/2}} \cdot \left(E[Y^3|N,W] - 3 \cdot E[Y^2|N,W] \cdot E[Y|N,W] + 2 \cdot E[Y|N,W]^3\right) \tag{3.6}$$

From equation 3.6,

$$E[Y^3|N,W] = (Sk[Y|N,W] \cdot V[Y|N,W]^{3/2}) + 3 \cdot (V[Y|N,W] \cdot E[Y|N,W]) + (E[Y|N,W])^3 \tag{3.7}$$

Moreover since,

$$E[Y^3|N] = \int E[Y^3|N,W] \cdot dF(W), \tag{3.8}$$

[13]Nitrogen application rates in Iowa typically range from 125 to 200 lbs./acre (Babcock and Hennessy, 1996). Moreover, the lowest recommended nitrogen rate for Iowa soil for a feasible nitrogen-corn price ratio based on the maximum return to nitrogen is 99 lbs./acre for SC and 141 lbs./acre for CC (Sawyer, 2016).

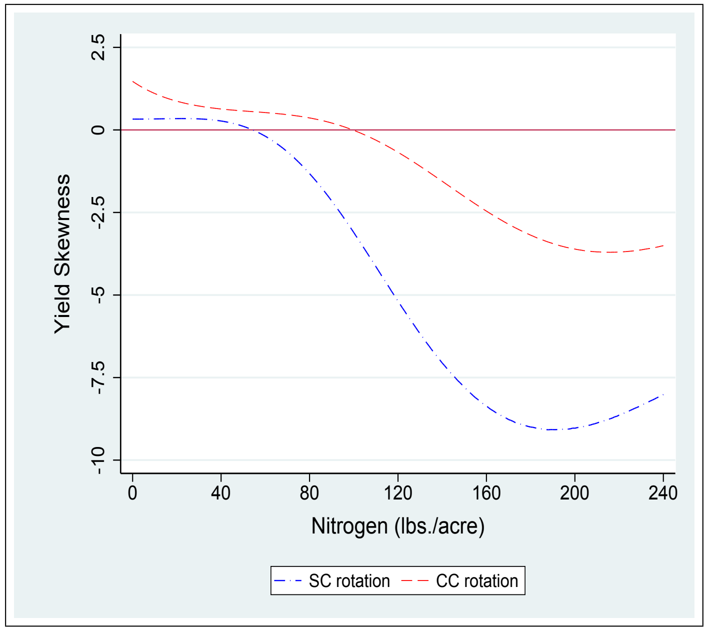



Figure 3.15 Yield Skewness conditional on nitrogen for the Ames farm under weather uncertainty.

Using equation 3.7 in 3.8, and 3.8 in 3.6, nitrogen conditional yield skewness under weather uncertainty is estimated.

Figure 3.15 plots the estimated yield skewness conditional on nitrogen under weather uncertainty (confidence intervals are not shown). This reconfirms the basic result that more nitrogen application increases the yield skewness (makes it more negative). It can be seen in figure 3.15 that increase in nitrogen increases yield skewness (i.e. makes it negative or more negative), which implies that the chances of realization of a higher than average yield subsequently increase.

## 3.6 Conclusion

This chapter uses a GLM with beta density to model the moments of the yield distribution conditional on nitrogen and weather variables (that includes growing season precipitation and TDD, and pre-season precipitation). The beta regression is a flexible representation of a stochastic production function which model the moments of the distribution of the output. The estimation of a separate precision sub-model in addition to the polynomials of nitrogen and the weather variables allow flexible estimation of the model.

This chapter outlines the effect of nitrogen and the weather variables on the first three moments of the yield distribution. Although, there has been previous research that has modeled the effect of weather variables on the moments of the yield distribution, this research work sets precedence by modeling the yield distribution conditional on nitrogen and weather variables. Moreover, it also discusses the effect of weather variables on the marginal product of nitrogen and its implication for the choice of optimal nitrogen.

The mean yield is non-decreasing in nitrogen, indicating a non-negative MPN. There is evidence of diminishing MPN, which is possibly convex in nitrogen. Extreme weather stress caused by hot and dry weather substantially reduces the mean MPN and the mean yield. Yield variability and yield skewness is estimated through integrating back the weather uncertainty into the yield variance and yield skewness estimates of nitrogen and weather conditional yield density. There is no evidence of yield variance to be increasing in nitrogen. Estimates of skewness conform to results of previous studies that at low level of nitrogen, yield skewness is positive and it eventually becomes negative as the amount of nitrogen applied increase. Yield skewness is negative in the range of nitrogen levels that are relevant for farming.

The main results of this chapter are used to define objective estimates of yield distribution, which includes the the mean yield and the MPN. These two estimates are used in the next chapter as a benchmark for comparison to subjective estimates of yield and MPN conditional on nitrogen. Moreover, the results of negative yield skewness at high levels of nitrogen application, is also used qualitatively (in terms of the sign of the skew) for comparing the subjective estimates of yield skewness.

## 3.7 Appendix

Table 3.4 SC Yields: Summary Statistics

| Nitrogen (lbs./acre) | N | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| 0 | 45 | 115.32 | 23.29 | 62.80 | 173.16 |
| 40 | 31 | 142.23 | 27.04 | 68.30 | 190.78 |
| 60 | 14 | 176.19 | 20.64 | 131.24 | 215.71 |
| 80 | 31 | 162.46 | 30.11 | 66.59 | 208.62 |
| 120 | 45 | 181.70 | 31.97 | 67.63 | 237.23 |
| 160 | 31 | 183.32 | 35.73 | 62.47 | 244.16 |
| 180 | 14 | 202.80 | 25.60 | 155.37 | 241.95 |
| 200 | 31 | 181.32 | 38.41 | 65.41 | 243.52 |
| 240 | 45 | 191.19 | 34.57 | 70.01 | 247.78 |

Table 3.5 CC Yields: Summary Statistics

| Nitrogen (lbs./acre) | N | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| 0 | 46 | 70.15 | 20.15 | 39.80 | 142.65 |
| 40 | 31 | 100.71 | 24.45 | 56.17 | 178.62 |
| 60 | 15 | 132.02 | 22.59 | 86.20 | 173.05 |
| 80 | 31 | 123.34 | 25.10 | 64.05 | 181.70 |
| 120 | 46 | 151.86 | 31.40 | 63.18 | 207.51 |
| 160 | 31 | 151.72 | 32.21 | 67.84 | 210.79 |
| 180 | 15 | 178.47 | 31.55 | 119.60 | 227.41 |
| 200 | 31 | 155.00 | 36.31 | 63.37 | 224.32 |
| 240 | 46 | 167.13 | 36.99 | 64.72 | 232.58 |

Table 3.6 Parameter estimates of GLM with beta denisty

| | (1) Yield | (2) Yield | (3) Yield |
|---|---|---|---|
| [1em] *Constant* | 0.90*** | 0.71*** | 0.92*** |
| | (0.55e-1) | (0.64e-1) | (0.53e-1) |
| *Sutherland* | -0.27e-1 | -0.43*** | -0.42*** |
| | (0.60e-1) | (0.57e-1) | (0.63e-1) |
| *Kanawha* | -0.79*** | -0.96*** | -1.04*** |
| | (0.61e-1) | (0.80e-1) | (0.79e-1) |
| *Nashua* | -0.44*** | -0.56*** | -0.70*** |
| | (0.58e-1) | (0.76e-1) | (0.62e-1) |
| *Rotation* | -0.71*** | -0.60*** | -0.70*** |
| | (0.37e-1) | (0.51e-1) | (0.43e-1) |
| *Rotation* × *Sutherland* | | | 0.18** |
| | | | (0.71e-1) |
| *Rotation* × *Kanawha* | | | -0.13** |
| | | | (0.64e-1) |
| *Rotation* × *(Kanawha/Nashua)* | | -0.23*** | |
| | | (0.65e-1) | |
| *Trend* | 0.36*** | 0.51*** | 0.44*** |
| | (0.58e-1) | (0.56e-1) | (0.53e-1) |
| *Trend* $^2$ | | -0.12** | |
| | | (0.56e-1) | |
| *Precipitation* | 0.26*** | -0.53e-1 | 0.14** |
| | (0.53e-1) | (0.66e-1) | (0.64e-1) |
| *Precipitation* $^2$ | 0.20*** | | 0.16*** |
| | (0.40e-1) | | (0.38e-1) |
| *Degree Days* | | -0.33e-1 | -0.21*** |
| | | (0.65e-1) | (0.60e-1) |
| *Degree Days* $^2$ | | -0.88e-1** | -0.17*** |
| | | (0.42e-1) | (0.28e-1) |
| *Preseason Precipitation* | -0.54e-3 | | 0.99e-1 |
| | (0.67e-1) | | (0.68e-1) |
| *Precipitation* × *Degree Days* | | 0.43*** | 0.21** |
| | | (0.16) | (0.11) |
| *Precipitation* × *Preseason Precipitation* | -1.66*** | | -1.70*** |
| | (0.20) | | (0.22) |

Table 3.6 (Continued)

| | | | |
|---|---|---|---|
| *Nitrogen* | 0.88*** | 0.91*** | 0.80*** |
| | (0.47e-1) | (0.46e-1) | (0.38e-1) |
| *Nitrogen* $^2$ | -0.29*** | -0.32*** | -0.31*** |
| | (0.21e-1) | (0.0221) | (0.0259) |
| *Nitrogen* $^3$ | 0.52e-1*** | 0.57e-1*** | 0.60e-1*** |
| | (0.19e-1) | (0.19e-1) | (0.16e-1) |
| *Nitrogen* × *(Sutherland/Kanawha)* | -0.30*** | -0.35*** | -0.28*** |
| | (0.55e-1) | (0.53e-1) | (0.46e-1) |
| *Nitrogen* × *Rotation* | 0.22*** | 0.23*** | 0.25*** |
| | (0.49e-1) | (0.53e-1) | (0.41e-1) |
| *Nitrogen* × *Trend* | 0.33*** | 0.33*** | 0.29*** |
| | (0.69e-1) | (0.62e-1) | (0.56e-1) |
| *Nitrogen* × *Precipitation* | 0.36*** | 0.47*** | 0.33*** |
| | (0.73e-1) | (0.71e-1) | (0.55e-1) |
| *Nitrogen* × *Degree Days* | | -0.96e-1 | -0.106** |
| | | (0.49e-1) | (0.43e-1) |
| *Nitrogen* × *Preseason Precipitation* | | | 0.22*** |
| | | | (0.77e-1) |
| *Nitrogen* × *Degree Days* $^2$ | | | -0.69e-2 |
| | | | (0.34e-1) |
| *Nitrogen* $^2$ × *Rotation* | | | -0.81e-1** |
| | | | (0.34e-1) |
| *Nitrogen* $^2$ × *Degree Days* | | | -0.38e-1 |
| | | | (0.30e-1) |
| *Nitrogen* $^2$ × *Degree Days* $^2$ | | | -0.93e-1*** |
| | | | (0.24e-1) |

| **Scale** | (1) | (2) | (3) |
|---|---|---|---|
| *Constant* | 3.57*** | 3.18*** | 4.06*** |
| | (0.24) | (0.20) | (0.22) |
| *Sutherland* | | | 0.28e-1 |
| | | | (0.26) |
| *Kanawha* | 1.35*** | 1.28*** | 1.81*** |
| | (0.22) | (0.20) | (0.20) |
| *Nashua* | | 0.62*** | 0.51** |
| | | (0.20) | (0.22) |
| *Rotation* | 0.29** | | -0.46*** |
| | (0.13) | | (0.17) |
| *Rotation* × *Sutherland* | | | 0.68** |
| | | | (0.27) |
| *Rotation* × *Kanawha* | -1.01*** | | |
| | (0.30) | | |
| *Rotation* × *Nashua* | | | 1.17*** |
| | | | (0.27) |

Table 3.6 (Continued)

| | | | |
|---|---|---|---|
| *Trend* | 1.14*** | | |
| | (0.196) | | |
| *Trend* $^2$ | -0.38*** | | |
| | (0.15) | | |
| *Precipitation* | 1.48*** | 1.26*** | 1.42*** |
| | (0.24) | (0.21) | (0.24) |
| *Precipitation* $^2$ | | -0.58*** | |
| | | (0.16) | |
| *Degree Days* | | 0.43 | 1.23*** |
| | | (0.30) | (0.16) |
| *Degree Days* $^2$ | | 0.76*** | 1.04*** |
| | | (0.22) | (0.17) |
| *Preseason Precipitation* | -1.21*** | | -0.86*** |
| | (0.39) | | (0.32) |
| *Preseason Precipitation* $^2$ | 0.71*** | | |
| | (0.26) | | |
| *Precipitation* × *Degree Days* | | -1.86*** | |
| | | (0.60) | |
| *Precipitation* × *Preseason Precipitation* | -2.87*** | | -3.51*** |
| | (0.75) | | (0.88) |
| *Nitrogen* | -0.12 | | -0.39*** |
| | (0.96e-1) | | (0.11) |
| *Nitrogen* $^2$ | | | -0.93e-1 |
| | | | (0.78e-1) |
| *Nitrogen* × *Nashua* | -0.70*** | | |
| | (0.22) | | |
| *Nitrogen* × *Rotation* | | | 0.58*** |
| | | | (0.19) |
| *Nitrogen* × *Rotation* × *Nashua* | | | -1.547*** |
| | | | (0.28) |
| *Nitrogen* × *Degree Days* | | | -0.51*** |
| | | | (0.18) |
| *Nitrogen* $^2$ × *Nashua* | | | 0.63*** |
| | | | (0.15) |
| *N* | 579 | 579 | 579 |
| *BIC* | -960.6 | -946.2 | -969.9 |

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

# CHAPTER 4. UNFOLDING THE BIAS IN FARM NUTRIENT MANAGEMENT

## 4.1 Introduction

This chapter studies the subjective beliefs of farmers surrounding their choice of optimal nitrogen and nitrogen management practices in their field. The primary goal of the chapter is to compare the subjective beliefs of the farmers with an objective agronomic benchmark. The objective benchmark is developed in chapter 3 of the dissertation, which is representative of actual (agronomic) relationship between yield distribution and nitrogen. This chapter investigate the subjective beliefs surrounding nutrient management, and report any significant differences with the objective benchmark, if it exist. While said so, it is important to note that providing a rationale for the choice of optimal nitrogen is beyond the scope of this research. This chapter rather provides factual and empirical evidence of the subjective expectations of farmer around their chosen level of nitrogen.

The results in chapter 3 and previous research studies have shown that nitrogen application alters the crop yield distribution. Researchers have found that since nitrogen application and its interaction with other inputs affect the moments of the yield distribution, nitrogen use in agricultural production is not limited to increasing the productivity of a field, but can also be used as a tool for risk management. A farmer undertakes a nutrient decision in his field in an uncertain environment. The nutrient decision making under uncertainty is a complex decision for a farmer, which is influenced by several factors, and also influence those factors, for example agricultural insurance, farm management practice, environmental externalities, technology adoption etc. Therefore, nutrient decision is at the center of agricultural decision -making, which has its relevance for several aspects of agricultural productivity, risk management and agricultural policy making.

The nitrogen decision made by a farmer hinges on their subjective belief about the nitrogen yield relationship, coupled with their perception of uncertainty and risk preferences over the uncertain environment. Agronomists and agricultural economists have studied the optimal fertilizer application rate for the farmers on their fields with an implicit assumption that farmers have rational expectations regarding the yield nitrogen beliefs and their perception of uncertainty. Most studies have commonly used data on yield ex-post the growing season from nitrogen trial experiments (Cerrato and Blackmer, 1990; Sawyer et al., 2015; Rajsic and Weersink, 2008) or county and farm level data (Horowitz and Lichtenberg, 1993; Smith and Goodwin, 1996), which are used to build rational expectations of farmers. In the light of these constructed rational expectations, many studies have analyzed the optimal nitrogen choice and have concluded that farmers over apply nitrogen on their fields (where an over application of nitrogen is defined as the nitrogen applied in excess of the agronomic recommended rate (Sheriff, 2005)).

This chapter presents the main results of the dissertation which builds upon the data and findings of previous chapters. The chapter is organized as follows. Section 4.2 discusses the scarce but relevant literature that exists in this context. Section 4.3 describes the multilevel linear regression model for subjective yield expectations and the variables used. The section also guide the readers through the development of the multilevel model, starting with an OLS linear regression model. Section 4.4 discuss in details the estimation of the multilevel model. The results are presented in three subsections of the results section. While the first subsection compares the subjective yield and marginal productivity of nitrogen (MPN) estimates of farmers with the objective yield and MPN estimates developed in chapter 3, subsection two describes the association of the subjective yield expectations and the MPN estimates with farmer and field specific variables (contextual effects). The third subsection of the results measure the subjective yield skewness and make a general comparison to the yield skewness results in chapter 3 and previous studies. Section 4.5 discusses the implication of the results and concludes.

## 4.2 Literature Review

The evaluation of optimal nitrogen application has been mostly based on ex-post outcomes (with the exception of SriRamaratnam et al. (1987)). The ex-post realized yield outcome assumes away uncertainty that is inherent in the nutrient decision making process faced by the farmers. SriRamaratnam et al. (1987) have measured producer yield beliefs through a survey eliciting expectations about mean yield and variance for three different levels of nitrogen. Their results supported that farmers overestimate the yield response to nitrogen. In their survey, yield and price expectations were elicited using a probability interval method, where respondents were asked to provide numeric weights (probabilities) to rank pre-specified intervals that were consistent with their subjective degree of belief. Respondents were asked to revise their beliefs subsequent to the presentation of results from the nitrogen yield experiments. Revision in their yield beliefs were not significant, and occurred only for extreme level of fertilizer. The results about the yield variance also indicated that the majority farmers believed nitrogen to be a risk reducing input.

Cerrato and Blackmer (1990) and Babcock (1992) have argued that if a farmer's perception of the yield nitrogen relationship is quadratic and increasing in nitrogen, the optimal nitrogen application rates are much higher relative to the von Liebig functional form, which can be a reason that drives excess use of nitrogen. Babcock and Blackmer (1992) and Babcock (1992) have stated that nitrogen acts as an insurance in the stochastic production process. The research mentions that the MPN is positive and diminishing, and if the MPN is also convex in nitrogen, a decision rule of thumb to apply more than the required amount of nitrogen is consistent with the weather uncertainty, and the uncertainty about the soil nitrogen availability. The average gain relative to the average loss of an additional unit of nitrogen is higher in situations, where nitrogen is the limiting factor compared to a situation of nitrogen abundance. Moreover, Babcock (1992) states that perceived over-application of nitrogen can also result if the mean of the subjective distributions of the MPN are much higher than the mean MPN of yield nitrogen relationships used to define nitrogen over-application.

Many research studies have found amount of nitrogen application to be correlated with several other farm variables like land tenancy, source of nitrogen advice, farm characteristics etc. Preckel et al. (2000) have found that the prevalence of tournament contracts[1] can drive nitrogen over application in the seed corn sector. Since, the land contracts are renewed based on yield history, insecure contracts encourage greater application of fertilizer inputs relative to certain contracts. Other studies that have found empirical relationship between nitrogen application and land tenancy arrangements include Horowitz and Lichtenberg (1993) and Smith and Goodwin (1996). While Horowitz and Lichtenberg (1993) expected a negative correlation, they found leased acres to be associated with higher nitrogen application. Smith and Goodwin (1996) reported lesser nitrogen usage per acre among farmers who have higher percentage of rented acres. Lawley et al. (2009) have reported that fertilizer dealers and independent crop consultants are likely to recommend increase in fertilization rates, while farmers who make their own nutrient plan or seek extension's recommendation are more likely to reduce fertilization rates. Moreover, the results of the study also showed that larger farms and land with greater slope are more likely to adopt an efficient nutrient management plan. Mishra et al. (2005) have found that farmers apply less fertilizer in the presence of highly erodible land, which they argued might be due to farmer's belief of lower marginal productivity of fertilizer in farms with highly erodible land. Chang and Mishra (2012) have found off-farm income to be inversely related to chemical usage in the presence of crop insurance. This was explained by income diversification and risk reduction through more opportunities available off-farm.

## 4.3 Data and Model

The data collected from the survey of subjective expectations as described in chapter 2 is used for the empirical analysis of subjective beliefs about nitrogen. Moreover, the results from chapter 3, which includes the modeled yield distribution is used as a comparable benchmark

[1] Tournament contracts are short term contracts, common in agriculture which are comprised of a fixed payment plus a bonus if the realized yield exceeds the average yield under the contract or a penalty if the realized yield falls short of the average yield under the contract.

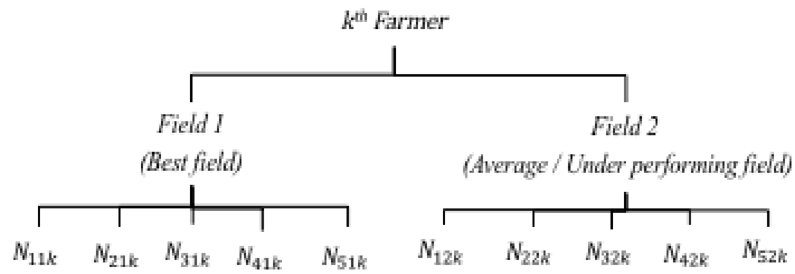



Figure 4.1 Hierarchical representation of subjective beliefs data

for the estimated subjective belief. Henceforth, the benchmark model of chapter 3 is referred to as the objective yield model.

The sample of farmer respondents came from the members of a co-operative in central Iowa. Since the research goal of the study is to investigate the subjective beliefs of farmers surrounding the choice of nitrogen decision and farm practices under uncertainty, the co-operative was instructed to send invitations to its member who have a purchase history of some minimal amount of fertilizer from the co-operative in the past. Therefore, the sample of respondents cannot be claimed to be randomly drawn, nor can be thought to be representative of farmers population in Iowa farmers. However, the unrepresentative feature of the sample do not reduce the significance of the results, as the primary research question focuses on the individual nitrogen beliefs of farmers while results associated with nitrogen beliefs on an average farmer is an ancillary research question.

### 4.3.1 Data description

The description of the survey methodology in chapter 2 indicates that the data collected for the sample respondents produce repeated measures data, where the repetition is over the levels of nitrogen nested in two different fields for every $k^{th}$ farmer. The hierarchy (nesting) in the survey data is represented in figure 4.1.

It is assumed that the respondents in the sample of farmers are independent. Farmers are at the highest level (or level 1) of the unit of analysis, indexed by $k$. Every $k^{th}$ farmer is asked to record their response for two fields. The first field is primed as the best performing field, and the priming of the second field is randomized across an average producing or an under performing field. The second level corresponds to the field, denoted by $j = \{1, 2, 3\}$, such that field 1, 2 and 3 are the best producing, average producing and under-performing field respectively. Within each field, monthly nitrogen application in lbs./acre are elicited. Moreover, conditional on the reported nitrogen application schedule, expected yield (in bu./acre of corn) are elicited from the responding farmer. The sum of reported monthly nitrogen application is denoted by $N_{3jk}$, and expected yield corresponding to the nitrogen application is denoted by $Y_{3jk}$. Moreover, based on their most recent (last month's) nitrogen application, farmer respondents are asked to report the yields corresponding to four counterfactual nitrogen application around $N_{3jk}$.[2] Every farmer is asked to record their response corresponding to two fields at the level 2 (field level), which produces a repeated measures data for a farmer across two fields. At the field level, for every farmer five yield nitrogen data points are recorded for the specific field. This produces repeated data measurement at the nitrogen level (or level 3) within any single field across farmer respondents. It can be seen in figure 4.1, $N_{ijk}$ denotes the counterfactual nitrogen application for $i = \{1, 2, 4, 5\}$ around $i = \{3\}$, which is the actual nitrogen application in the $j^{th}$ field of the $k^{th}$ farmer. $Y_{ijk}$ is the yield that $k^{th}$ farmer expects his $j^{th}$ field to produce corresponding to $N_{ijk}$ amount of nitrogen application.

As already mentioned earlier, the primary goal of this study is to investigate the subjective beliefs of individual farmers surrounding their nitrogen management decision. The marginal effect of nitrogen within a field for any farmer and across farmers average effect are illustrated in figure 4.2. Only few selected farmers and their best field are included in figure 4.2 in order to juxtapose the average and marginal effects of nitrogen on the expected yield. As it can be seen in figure 4.2, the red hollow circles plot the reported amount of (chosen) actual nitrogen application (in lbs./acre), and the corresponding expected yield (in bu./acre) across farmers $((N_{3jk}, Y_{3jk}))$. The red solid line fits a linear slope across farmers' subjective yield and

---

[2] The four counterfactual nitrogen application treatment comprised of 75%, 85%, 110% and 125% of their actual nitrogen application reported in the last month. Their subjective belief about nitrogen yield response is measured through the four counterfactual nitrogen applications around their actual (or planned) nitrogen application denoted by $N_{3jk}$.

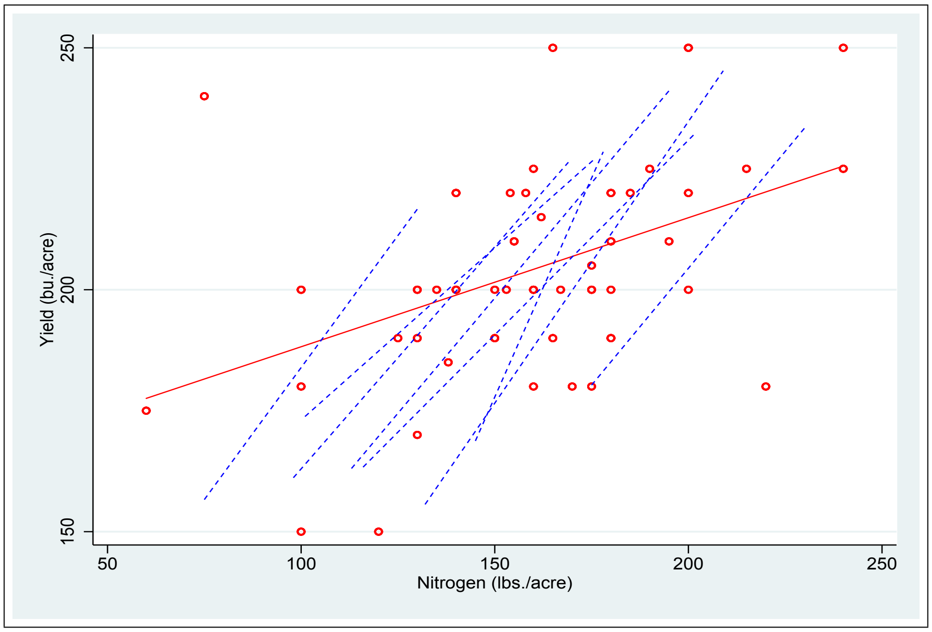



Figure 4.2 Illustrating average and individual subjective beliefs about nitrogen yield relationship. Each red hollow circle indicates the chosen optimal nitrogen level by a farmer in his specific field, and the corresponding expected yield. The solid red line depicts the average nitrogen yield response across farmers, which is the linear regression of elicited expected yield across all farmers at their chosen nitrogen levels (i.e. the linear fit across red hollow circles). The dashed blue lines (shown only for selected eight farmers) measure the individual subjective beliefs about nitrogen yield response for the selected farmers on their respective field, which is the linear fit of the expected yield across the five different level of nitrogen values (not shown here) corresponding to which the expected yield are elicited.

their (chosen) actual nitrogen application values ($(N_{3jk}, Y_{3jk})$ denoted by red hollow circles) reported by farmers. The blue dashed lines are the linear fit depicting the subjective nitrogen yield response belief for the $k^{th}$ farmer in their best field across the five nitrogen application levels (actual and counterfactual denoted by $(N_{ijk}, Y_{ijk})$ for $i = \{1, 2, 3, 4, 5\}$ ). The subjective belief of nitrogen yield response denoted by the slope of the blue dashed line is a measure of the subjective belief about MPN, which is of primary research interest in this study.

The description of the survey methodology and the data structure makes it apparent that the yield responses are not independent across observations. It is assumed that the nitrogen conditional yield response across farmers are independent. Therefore, for any two farmers $k$ and $k'$,

$$Cov(Y_{ijk}, Y_{i',j',k'}) = 0,\ \forall\, k \neq k',\ j, j' = \{1, 2, 3\},\ \forall\, i, i' = \{1, 2, 3, 4, 5\} \tag{4.1}$$

Unlike the independence of observations across farmers, yield responses reported by the same farmer respondent are likely to be correlated. There may be an unobserved correlation among the responses of the $k^{th}$ farmer across the two fields and within a field due to the farmer specific effect or a field specific effect or both. Therefore, for any farmer $k$,

$$Cov(Y_{ijk}, Y_{i',j',k}) \neq 0,\ \forall\, k,\ \forall\, (j, j') = \{(1, 2), (1, 3)\},\ (i, i') = \{1, 2, 3, 4, 5\} \tag{4.2}$$

$$Cov(Y_{ijk}, Y_{i',j,k}) \neq 0,\ \forall\, k,\ \forall\, j = \{(1, 2, 3)\},\ \forall\ \ i \neq i',\ i, i' = \{1, 2, 3, 4, 5\} \tag{4.3}$$

In context to the modeling of subjective yields, three different regression methods are considered. The three models are discussed, and the one that can appropriately model the nested correlation structure among the data is chosen to present the results.

### 4.3.2 Individual farmer and field specific OLS linear regression

An individual farmer and field specific linear regression can be fit to the five points of the subjective yield nitrogen relationship denoted by $(N_{ijk}, Y_{ijk})$ for $\forall\ i = \{1, 2, 3, 4, 5\}$, for every $j^{th}$ field of the $k^{th}$ farmer. The regression model is described as below

$$Y_{ijk} = \alpha_{0,jk} + \alpha_{1,jk} \cdot N_{ijk} + \nu_{ijk},\ \ \forall\, j = \{1, 2, 3\},\ k = \{1, 2.., K\} \tag{4.4}$$

where $K$ is the total number of farmers. $N_{ijk}$ is the amount of nitrogen for $i = \{1, 2, 3, 4, 5\}$ and $\nu_{ijk}$ is the error term. $\alpha_{0,jk}$ is the intercept term for the regression equation for $j^{th}$ field of the $k^{th}$ farmer. Similarly, $\alpha_{1,jk}$ is the linear slope of the nitrogen term $N_{ijk}$ for the regression equation for $j^{th}$ field of the $k^{th}$ farmer. Under the regression model denoted by equation 4.4, $\alpha_{1,jk}$ is the estimate of subjective MPN for $j^{th}$ field of the $k^{th}$ farmer.

For every $(j, k)$ combination the individual regressions have three degrees of freedom, since there are five observations (corresponding to five nitrogen application levels). With just five observation points, in an individual farmer and field specific regression, including additional regressors is not a choice. Hence, correlation of subjective yields with other farm or farmer specific variables cannot be estimated. Moreover, estimating an individual farmer and field specific regression ignores the correlation that exists in the data, which can be used to build a better informed model if the correlation structure in the data is appropriately accounted for.

### 4.3.3 Pooled OLS regression

An alternative to individual farmer and field specific OLS regression is a pooled OLS regression. Pooled OLS regression uses stacked data across farmers, fields and nitrogen levels so that each observation is treated to be an individual unit. In a pooled regression, the dependence of the observations in the data structure is ignored, and it is assumed that the sample observations are independent. The individual fixed effects at the farmer and the field level are captured by allowing for a farmer and field specific fixed effect. Additionally, farmer and field specific slope for nitrogen are also introduced. The model is described as,

$$\begin{aligned} Y_{ijk} = {} & \Sigma_{k=1}^{K} \left(\alpha_{0,1k} \cdot D_{1k} + \alpha_{0,2k} \cdot D_{2k} + \alpha_{0,3k} \cdot D_{3k}\right) + \Sigma_{k=1}^{K} \alpha_{1,1k} \cdot (N_{ijk} \times D_{1k}) \\ & + \Sigma_{k=1}^{K} \alpha_{1,2k} \cdot (N_{ijk} \times D_{2k}) + \Sigma_{k=1}^{K} \alpha_{1,3k} \cdot (N_{ijk} \times D_{3k}) + \alpha_{\mathbf{2}} \cdot \mathbf{S_k} + \alpha_{\mathbf{3}} \cdot \mathbf{Z_{jk}} \\ & + \nu_{ijk}, \qquad \forall\, j = \{1, 2, 3, \},\ k = \{1, 2.., K\} \end{aligned} \tag{4.5}$$

$D_{1k}$, $D_{2k}$ and $D_{3k}$ are the respective dummy variables for the best producing, average producing and under-performing fields of the $k^{th}$ farmer. $\mathbf{S_k}$ are the (level 1) variables at the level of farmers, $\mathbf{Z_{jk}}$ are the farmer and field specific variables at (level 2), and $N_{ijk}$ is the (level 3 variable) nitrogen application levels in $j^{th}$ field of the $k^{th}$ farmer. A separate farmer and field specific intercept for the best and average or under-performing field of the $k^{th}$ farmer is denoted by $\alpha_{0,1k}$ and $\alpha_{0,2k}$ or $\alpha_{0,3k}$ respectively. Similarly, $\alpha_{1,1k}$ and $\alpha_{1,2k}$ or $\alpha_{1,3k}$ respectively denotes the farmer and field specific slope of the nitrogen variable, $N_{ijk}$ for the best and average or under-performing field of the $k^{th}$ farmer.

Although, a farmer and field specific coefficient for higher order polynomials of $N_{ijk}$ cannot be included (since, only 5 data points exist for each field of every farmer), higher order polynomials of nitrogen with a common coefficient across all farmers for every field can be included under the assumption that higher order effects of nitrogen application are same across all farmers and fields.

Alternatively, higher order nitrogen terms across all farmers specific to a field type can be included, which assumes that the effect of higher order nitrogen polynomials are same across farmers for the same field type, but differ across type of fields. While the first alternative regression with higher order polynomials (square and cube) of nitrogen will include additional regressor variables $N_{ijk}^2$ and $N_{ijk}^3$, the second alternative regression with higher order polynomial terms will have $(N_{ijk}^2 \times D_{1k})$, $(N_{ijk}^2 \times D_{2k})$ and $(N_{ijk}^2 \times D_{3k})$ for the quadratic terms, and $(N_{ijk}^3 \times D_{1k})$, $(N_{ijk}^3 \times D_{2k})$ and $(N_{ijk}^3 \times D_{3k})$ for the cubic terms. Whether or not to include the higher order polynomials depends on the degrees of freedom that the sample size allows. However, it is made under the assumption that the higher order nitrogen effects are same across farmers. Interaction terms of $\mathbf{S_k}$, $\mathbf{Z_{jk}}$ and $N_{ijk}$ can also be included in the model.

Although this may look like a desirable model, this require large number of parameters to be estimated, as it includes both farmer and field specific intercepts and slope. The pooled OLS model, implicitly assumes that the observations are independent. Although, the introduction of fixed effect through farmer and field specific intercept and nitrogen slope captures the individual farmer and field specific beliefs, this does not capture the unobserved correlation among the error terms. It assumes

$$Cov(v_{ijk}, v_{i',j',k'}) = 0,\ \forall\ (k,k') = \{1,..,K\},\ \forall\ (j,j') = \{(1,2),(1,3)\},\ \forall\ i,i' = \{1,2,3,4,5\}$$

This is not necessarily true except when $k \neq k'$ (by assumption). The error terms are the unobserved part of the yield expectations that cannot be explained by the observed variables or even farmer and field specific fixed effects. Therefore, it likely that unobserved commonality across the observations of the same units (farmers or the fields) is still present, which make the residuals (error terms) to be correlated. A unit in this context is a farmer or a field for which repeated measures at any level exist.

### 4.3.4 Multilevel regression model

The regression methods described in the previous section briefly explain the implications of the assumptions involved in each method, and why they are unfit for modeling the collected subjective yield expectations data. While comparing the pooled OLS regression model with the individual farmer and field specific OLS regressions, a significant difference exists in the underlying assumptions about the correlation structure in the data. The individual farmer and field specific OLS regressions do not acknowledge that the information in the observations across the same units can be pooled to use the correlation structure in the data. Pooled data gathers information about the higher level observed explanatory variables, which are constant for any given unit, but has variability across units. For example farmers' beliefs about nitrogen may be different for a farmer farming on leased land compared to a farmer farming on owned land. Although, the land ownership across nitrogen treatment for any given unit is same, but pooling the data can use information in the responses across farmers for differences in the land ownership. On the other hand, the pooled OLS regression ignores the correlation among the observations of the same unit by treating them to be independent. Therefore, while the individual OLS regression model do not fully exploit the information contained in the data, the assumptions of the pooled OLS model uses more information by treating every observation to be independent. For example, if across the 5 nitrogen treatment for any given field, the field CSR is 80, pooled OLS assumes that there are 5 fields, each with a CSR of 80, while the fact is that there is one single field.

A multilevel linear model (Raudenbush and Bryk, 2002; Snijders and Bosker, 2011) is used to model the subjective yields. A multilevel linear model consists of fixed and random components at different levels, based on the nesting structure of groups in the data. The regression coefficient of an explanatory variable of interest is modeled as a random variable which follows a pre-specified distribution. The mean of the distribution of the random regression coefficient is the fixed component of the regression model. The fixed component in the multilevel model should not be confused with the fixed effects used in the pooled OLS regression. While the fixed effects in the pooled OLS regression captures the individual unit specific intercept and slope, the fixed component in the multilevel model indicates the deterministic part of the regression coefficient of an explanatory variable. The deterministic part of the regression coefficients (which also includes those explanatory variables which do not have a random coefficient) are the fixed component of the model, which are common to all the units in the data. The heterogeneity in the data and appropriate correlation structure is accounted through the introduction of random coefficients. The random coefficients significantly reduces the number of parameters to be estimated and at the same time also accounts for the unobserved unit specific effects, which in a way also controls for clustering of data at various levels. However, the multilevel model imposes structure on the random effects that they are drawn from a parametric distribution (mostly from a normal distribution), which is not an unreasonable assumption to make when much is not known about the distribution of the sample from theory. Moreover, practically it is not possible to include as many random effects as desired. Therefore, it is assumed that effects of variable without random components specification is same across all units,. This is again not an unreasonable assumption if the choice of variable for including random effects is guided by theory or the research question.

The use of multilevel model also allows identification of contextual effects in the regression model. In cognitive psychology, contextual effects are defined as the influence that the environment or the surroundings of an individual (or a group level variable) have on the effect of the individual unit independent variable on the dependent variable (Diez, 2002). For example in the subjective yield model, the perception about the effect of nitrogen on the yield is the integral part of the model, but a contextual effect may be defined as how does the perception differ across farmers who are more educated or who have adopted a delayed planting date, as these are the environment (context) of the farmer under which he has reported his beliefs.

A multilevel model is formed by specifying the regression equation at the lowest level, and the regression coefficients of the lowest level regression equation are functions of higher level explanatory variables. A three level hierarchical model is generally described as in the following equations:

$$Level\ 3: \quad Y_{ijk} = \gamma_{0jk} + \gamma_{1k} \cdot X_{ijk} + \epsilon_{ijk} \tag{4.6}$$

$$Level\ 2: \quad \gamma_{0jk} = \phi_{00k} + \phi_{01k} \cdot G_{jk} + \tau_{0jk} \ ; \qquad \gamma_{1jk} = \phi_{10k} + \phi_{11k} \cdot G_{jk} + \tau_{1jk} \tag{4.7}$$

$$Level\ 1: \quad \phi_{00k} = \theta_{000} + \theta_{001} \cdot D_k + \eta_{00k} \ ; \qquad \phi_{01k} = \theta_{010} + \theta_{011} \cdot D_k \tag{4.8}$$

$$\phi_{10k} = \theta_{100} + \theta_{101} \cdot D_k + \eta_{10k} \ ; \qquad \phi_{11k} = \theta_{110} + \theta_{111} \cdot D_k \tag{4.9}$$

The above equations represent a general form of the three level hierarchical linear model, where $X_{ijk}$ is the vector of level 3 explanatory variable and $Y_{ijk}$ is the dependent variable. $\epsilon_{ijk}$ is the error term at the level 3. It can be seen that the coefficient of $X_{ijk}$ is a function of level 2 set of variables denoted by $G_{jk}$ and the level 2 residuals (random components), $\tau_{0jk}$ and $\tau_{1jk}$. Similarly, the parameter coefficients of the level 2 explanatory variables are function of level 1 variables and the level 1 residuals. Substituting for the level 2 and level 1 equations in equation 4.6, the full multilevel model can be written as

$$\begin{aligned} Y_{ijk} =& \beta_0 + \beta_1 \cdot D_k + \beta_2 \cdot G_{jk} + \beta_3 \cdot X_{ijk} + \beta_4 \cdot (G_{jk} \times D_k) + \beta_5 \cdot (X_{ijk} \times D_k) \\ &+ \beta_6 \cdot (X_{ijk} \times G_{jk}) + \beta_7 \cdot (X_{ijk} \times G_{jk} \times D_k) + \tau_{0jk} + (\tau_{1jk} \times X_{ijk}) \\ &+ \eta_{00k} + (\eta_{10k} \times X_{ijk}) + \epsilon_{ijk} \end{aligned} \tag{4.10}$$

where the symbol '$\times$' denotes an interaction between two variables. $\tau_{0jk}$ and $\tau_{1jk}$ are the level 2 random slope and random intercept term whereas $\eta_{00k}$ and $\eta_{10k}$ are the level 1 random slopes and intercept. It is assumed that $\epsilon_{ijk} \sim N(\mathbf{0}, \sigma^2_\epsilon)$, $(\tau_{0jk}, \tau_{1jk}) \sim N(\mathbf{0}, \Gamma_\tau)$ and $(\eta_{00k}, \eta_{10k}) \sim N(\mathbf{0}, \Gamma_\eta)$, where $\Gamma_\tau$ and $\Gamma_\eta$ are the symmetric variance-covariance matrices of the random components, where

$$\Gamma_\tau = \begin{pmatrix} \sigma^2_{\tau 0} & \upsilon_{21} \\ \upsilon_{12} & \sigma^2_{\tau 1} \end{pmatrix} \qquad \Gamma_\eta = \begin{pmatrix} \sigma^2_{\eta 0} & \rho_{21} \\ \rho_{12} & \sigma^2_{\eta 1} \end{pmatrix}$$

### 4.3.5 Applying the multilevel regression model

Using the multilevel framework as described in the previous section, the subjective yields are modeled through the equation 4.11, which states the full reduced form equation of the level three multilevel model of subjective yield expectations as below:

$$Y_{ijk} = \beta_0 + \beta_{\mathbf{1}} \cdot \mathbf{\bar{D}_k} + \beta_{\mathbf{2}} \cdot \mathbf{\bar{G}_k} + \beta_{\mathbf{3}} \cdot \mathbf{\bar{G}_{jk}} + \beta_{\mathbf{4}} \cdot \mathbf{\bar{X}_{ijk}} + \beta_{\mathbf{5}} \cdot (\mathbf{\bar{D}_k} \times \mathbf{\bar{X}_{ijk}}) + \beta_{\mathbf{6}} \cdot (\mathbf{\bar{G}_k} \times \mathbf{\bar{X}_{ijk}})$$
$$+ \beta_{\mathbf{7}} \cdot (\mathbf{\bar{G}_{jk}} \times \mathbf{\bar{X}_{ijk}}) + \tau_{0jk} + (\tau_{\mathbf{1jk}} \times \mathbf{\bar{X}_{ijk}}) + \eta_{00k} + (\eta_{\mathbf{10k}} \times \mathbf{\bar{X}_{ijk}}) + \epsilon_{ijk} \qquad (4.11)$$

Table 4.1 describe the variables used to model the subjective yield expectations as in equation 4.11.

Table 4.1 Description of variables

| Variable | Description |
|---|---|
| **Level 3** | $\mathbf{\bar{X}_{ijk}}$ |
| $\bar{N}_{ijk}$ | *Nitrogen* (in lbs./acre) centered around chosen application level |
| $\bar{N}^2_{ijk}$ | Square of $\bar{N}_{ijk}$ |
| $\bar{N}^3_{ijk}$ | Cube of $\bar{N}_{ijk}$ |
| **Level 2** | $\mathbf{\bar{X}_{jk}}$ |
| $\bar{N}_{jk}$ | Choice of *nitrogen* centered around the (group) mean of chosen nitrogen level (farmer's both field) |
| **Level 2** | $\mathbf{\bar{G}_{jk}}$ |
| $(Average)_{jk}$ | Dummy variable for *Average* producing field |
| $(Under)_{jk}$ | Dummy variable for *Under* performing field |
| $(Fall)_{jk}$ | Dummy variable for most recent nitrogen application during *fall* |
| $(After\ Planting)_{jk}$ | Dummy variable for most recent nitrogen application *after planting* |
| $(Nitrogen\ Test)_{jk}$ | Dummy variable for soil or plant tissue *nitrogen test* |
| $(CSC\ rotation)_{jk}$ | Dummy variable for Corn-Soybean-Corn crop rotation |
| $(SCC\ rotation)_{jk}$ | Dummy variable for Soybean-Corn-Corn crop rotation |
| $(Mixed\ rotation)_{jk}$ | Dummy variable for mixed rotation (both corn and soybean) |
| $(Conventional)_{jk}$ | Dummy variable for *conventional* tillage |
| $(Planting)_{jk}$ | *Planting* date (in units of 15 days from April 15) |

Table 4.1 (Continued)

| Variable | Description |
|---|---|
| $(Pollination)_{jk}$ | Expected *pollination* date (in interval of 15 days from June 15 - 30) |
| $(C\bar{S}R)_{jk}$ | CSR centered around (group) mean CSR (farmer's both field) |
| $(Fiel\bar{d}\ Size)_{jk}$ | *Field size* centered around (group) mean field acres (farmer's both field) |
| $(Center\ Prob)_{jk}$ | Probability of yield lying within $\pm$ 10% of the expected yield |
| $(Skew\ Prob)_{jk}$ | Difference in cumulative yield probability in the left vs. the right tail (tail in $\pm$ 10% of the expected yield) |
| **Level 1** | $\mathbf{\bar{X}_k}$ |
| $\bar{N}_k$ | Farmer's (group) mean of chosen *nitrogen* levels centered around (grand) mean of *nitrogen* application across farmers |
| **Level 1** | $\mathbf{\bar{G}_k}$ |
| $(C\bar{S}R)_k$ | Farmer's (group) mean CSR centered around (grand) mean of *CSR* |
| $(Fiel\bar{d}\ Size)_k$ | Farmer's (group) mean *field size* centered around (grand) mean *field size* |
| **Level 1 variables** | $\mathbf{\bar{D}_k}$ |
| $(Shared\ decision)_k$ | Dummy variable for nutrient management *decision shared* |
| $(Education)_k$ | Dummy variable for farmer's *education* beyond high school |
| $(Expe\bar{r}ience)_k$ | Farming *experience* centered around (grand) mean of *experience* |
| $(Land\ \bar{f}armed)_k$ | *Land farmed* centered around (grand) mean of farmed acres |
| $(\%\ Land\ owned)_k$ | Percentage of farmed *land owned* by farmer |

The variables at all levels are centered to facilitate meaningful interpretation of the estimated parameters. A bar over a variable in table 4.1 indicates that the variable has been centered. The choice of centering is contextual and depends on the research questions that need to be answered (Hofmann and Gavin, 1998; Enders and Tofighi, 2007). Level 1 variables are centered around the grand mean, which is the average of the level 1 (i.e. across farmers) variables denoted by $\bar{\mathbf{D}} = \Sigma_k D_k$. The vector of level 1 centered variables is denoted by $\bar{\mathbf{D}}_{\mathbf{k}} = \Sigma_k D_k - \bar{\mathbf{D}}$. Level 2 variables, denoted by $\mathbf{G_{jk}}$ are centered at the group mean, which is the mean of the two fields of each farmer[3], and the group mean is centered at the grand mean, denoted by $\mathbf{G_k}$. The dummy variables at all levels are not centered. For variables that measure some probability or percentage of ownership of farmed land are also uncentered as heir value of zero has a natural interpretation. The planting dates are transformed by subtracting one, so that a value of zero indicates *planting* date around April 15, and increase in *planting* date by one unit indicates a delay of 15 days in planting. Similarly, one is also subtracted from *pollination* date, so that a value of zero indicates *pollination* date between June 15 - 30, and a unit increase shifts the expected pollination dates ahead by 15 days.

At the lowest level of the hierarchy in the variables is the nitrogen application. The amount of nitrogen applied at the level 3 is centered around the chosen (actual or planned) level of nitrogen application. This measures the marginal effect of nitrogen on the yield expectations around the chosen (or actual) level of nitrogen application by the farmers. At level 3, the nitrogen variable is centered around $N_{3jk}$, which is the chosen (actual or planned) level of nitrogen application by the $k^{th}$ farmer in their $j^{th}$ field. Since, nitrogen is at (lowest) level 3, it is also centered at level 2. At level 2, nitrogen variable is centered at the group mean, which is denoted by $\bar{N}_{jk}$. Therefore,

$$\bar{N}_{ijk} = N_{ijk} - N_{3jk}, \;\; \bar{N}_{jk} = N_{3jk} - \bar{N}_k, \;\; \bar{N}_k = \tilde{N}_k - \tilde{N},$$
$$where \;\; \tilde{N}_k = \frac{1}{n_{jk}} \Sigma_j N_{3jk} \;\; and \;\; \tilde{N}_k = \frac{1}{n_k} \Sigma_k N_k,$$

[3]group mean is same as the individual mean for farmers who have recorded response for one field only

and $n_{jk}$ and $n_k$ are the total number of fields for $k^{th}$ farmer and total number of farmers respectively).

Centering helps in identifying the between and within effects of a variable (here nitrogen application). The decomposition of the total effect in between and within effect distinguishes and tests whether the individual effects are significantly different from the average effects. While the between effects of nitrogen variable would be an average effect across farmers and their fields at the chosen application level of nitrogen (denoted by $\bar{N}_{jk}$), within effects represent the beliefs of farmers about the nitrogen productivity around their choice of nitrogen application in their field (denoted by $\bar{N}_{ijk}$). Within effects can be thought of as farmers' perception about the production function or the yield nitrogen mapping they perceive on their field, while a between effect can be conceptualized as an average effect of the optimal nitrogen choice on yield expectations across farmers.

Level 3 regression equation models the subjective yields corresponding to the $i^{th}$ level of nitrogen application on the $j^{th}$ field for the $k^{th}$ farmer, which regresses subjective yields on the polynomials of nitrogen. The coefficients of the intercept, linear and quadratic nitrogen terms at level 3 (except the cubic term) are random. This implicitly assumes that the third order polynomials have the common slope coefficient for all fields across any farmer[4]. This implies that all farmers believe that the curvature of subjective MPN is same at any given level of nitrogen application chosen. The intercept and the coefficient of the linear and quadratic nitrogen term are functions of level 2 field specific variables and random components. Similarly, the coefficients of level 2 fixed components are function of farmer specific level 1 variables and the random components at level 1.

The fixed components in the above equation are denoted with a coefficient $\beta$. The residual unexplained error term, $\epsilon_{ijk}$ are interpreted as the unobserved part of the farmer's expectation process to the researcher.

The random components in the equation 4.11 are denoted by the $\eta$ and $\tau$ terms. It must be noted that in comparison to the general framework of a three tier multilevel model in equa-

[4]Although a random term corresponding to cubic term of nitrogen was estimated, it could not be identified due to insufficient variation in the data.

tion 4.10, the random slopes included in the model correspond to nitrogen variables only. The perception of farmers about nitrogen effects on yield is of central interest in this study, hence the heterogeneity of the nitrogen effects across farmers is captured through the random slope parameters of nitrogen, which includes $\eta_{01k}$, $\eta_{10k}$ and $\eta_{20k}$ at the farmer level and $\tau_{01jk}$, $\tau_{1jk}$ and $\tau_{2jk}$ at the field level for the intercept, linear and quadratic nitrogen term. Therefore, the vector of random components at the farmer and the field level is given by $\eta = [\ \eta_{00k} \ \ \eta_{10k} \ \ \eta_{20k}\ ]'$ and $\tau = [\ \tau_{0jk} \ \ \tau_{1jk} \ \ \tau_{2jk}\ ]'$ respectively, such that $\eta \sim N(\mathbf{0}, \Sigma_\eta)$ and $\tau \sim N(\mathbf{0}, \Sigma_\tau)$, where $\Sigma_\eta$ and $\Sigma_\tau$ are defined as,

$$\Gamma_\eta = \begin{pmatrix} \sigma^2_{\eta 0} & \rho_{01} & \rho_{02} \\ \rho_{10} & \sigma^2_{\eta 1} & \rho_{12} \\ \rho_{20} & \rho_{21} & \sigma^2_{\eta 2} \end{pmatrix} \qquad \Gamma_\tau = \begin{pmatrix} \sigma^2_{\tau 0} & \upsilon_{01} & \upsilon_{02} \\ \upsilon_{10} & \sigma^2_{\tau 1} & \upsilon_{12} \\ \upsilon_{20} & \upsilon_{21} & \sigma^2_{\tau 2} \end{pmatrix}$$

## 4.4 Results

The model can be estimated using either the maximum likelihood estimation (MLE) or the restricted maximum likelihood estimation (REML). REML is chosen over MLE to estimate the model because REML takes into consideration the degrees of freedom used in the estimation of fixed components while estimating the variance component of the model. Moreover, in smaller sample REML estimation provides unbiased estimates of the parameters compared to MLE. (Snijders and Bosker, 2011).

### 4.4.1 Estimation of the random components of the regression

Four different multilevel models are estimated and reported in table 4.4 in appendix 4.6. Models 1 to 3 are based on the same random components matrix but different fixed effects. In model 2, variables capturing crop rotation are introduced. Moreover, in model 3 in addition to crop rotation, variables that measured the probability mass in the tails and in the center of the reported yield distribution are included. Since there are few farmer respondents, who have not elicited valid probability response (violated the probability axioms), they are excluded from model 3. In the first three models, model 1 to 3, only farmer specific random terms (at level 1) are introduced in the model, Therefore, the random component matrix to be estimated is $\Gamma_\eta$. In model 4, along with farmer level random effects, field level random effects are also introduced. Therefore, the random component matrix $\Gamma_\tau$ is also estimated along with $\Gamma_\eta$. It should be noted that in multilevel models although the heterogeneity of the individual units or groups and the covariance between them is captured through the random component matrix, the estimation involves estimating the variance-covariance matrix at the desired level. The individual random coefficients are rather predicted (and not estimated) using the shrinkage estimators (Snijders and Bosker, 2011). Therefore, in models one to three, and also in model 4, the residual variance in the subjective yields is decomposed into variability due to farmers. In model 4 the residual variability (after accounting for farmer random effects) is further decomposed into variability due to different fields. The details of groups at the level 1 (farmers) and level 2 (fields) is described in table 4.2 below.

Table 4.2 Summary of group variable

| | | No. of | No. of observations per group | | |
|---|---|---|---|---|---|
| Model | Group | Groups | Minimum | Average | Maximum |
| Model 1 | Farmers | 65 | 5 | 8.3 | 10 |
| Model 2 | Farmers | 65 | 5 | 8.3 | 10 |
| Model 3 | Farmers | 59 | 5 | 8.1 | 10 |
| Model 4 | Farmers | 65 | 5 | 8.3 | 10 |
| | Fields | 108 | 5 | 5 | 5 |

While estimating the random component matrix $\Sigma_\eta$ for models one to three, all terms of $\Sigma_\eta$ cannot be identified due to insufficient variation in the sample data. Therefore, $\rho_{20} = \rho_{21} = 0$ was assumed (i.e. the farmer specific random slope corresponding to quadratic term of nitrogen is uncorrelated to random slope corresponding to linear term of nitrogen or random intercept). Hence, the random component matrix to be estimated is $\Gamma_{\eta 1}$ such that

$$\Gamma_{\eta 1} = \begin{pmatrix} \sigma^2_{\eta 0} & \rho_{01} & 0 \\ \rho_{10} & \sigma^2_{\eta 1} & 0 \\ 0 & 0 & \sigma^2_{\eta 2} \end{pmatrix}$$

The likelihood ratio test is conducted to test the null hypothesis $\sigma^2_{\eta 2} = 0$. Likelihood ratio test for testing the variance component of the random parameters matrix is one sided test, since variance is non-negative. Hence, in order to test the one sided null hypothesis with valid degrees of freedom, Snijders and Bosker (2011) suggested to use a 50:50 mixture of chi-square distribution with zero and one degrees of freedom, to test and determine the critical rejection region for the null hypothesis that the variance of a random component is zero. The likelihood ratio test could not reject the null that $\sigma^2_2 = 0$, using the test statistic, which is a 50:50 mixture of chi-square distribution with zero and one degrees of freedom. Further tests rejected the null that $\sigma^2_{\eta 1} = 0$ or $\sigma^2_{\eta 0} = 0$ or the covariance between the random slope and random intercept is zero (i.e. $\rho_{10} = 0$). Therefore, the final random component matrix to be estimated for models 1 to 3 is $\tilde{\Gamma}_\eta$ such that

$$\tilde{\Gamma}_\eta = \begin{pmatrix} \sigma^2_{\eta 0} & \rho_{01} \\ \rho_{10} & \sigma^2_{\eta 1} \end{pmatrix}$$

For model 4, as already mentioned random effects at both the farmer and the field level are estimated. The two random component matrices to be identified are $\Sigma_\eta$ and $\Sigma_\tau$. For both matrices, covariance between the quadratic term of nitrogen with the random slope of linear nitrogen term and the random intercept term could not be identified and, hence was assumed to be zero (ie.$\rho_{20} = \rho_{21} = 0$ and $\upsilon_{20} = \upsilon_{21} = 0$ ). Therefore, the random component variance-covariance matrices to be estimated are given by $\Gamma_{\eta 1}$ and $\Gamma_{\tau 1}$ at the farmer and the field level respectively as below:

$$\Gamma_{\eta 1} = \begin{pmatrix} \sigma^2_{\eta 0} & \rho_{01} & 0 \\ \rho_{10} & \sigma^2_{\eta 1} & 0 \\ 0 & 0 & \sigma^2_{\eta 2} \end{pmatrix} \qquad \Gamma_{\tau 1} = \begin{pmatrix} \sigma^2_{\tau 0} & \upsilon_{01} & 0 \\ \upsilon_{10} & \sigma^2_{\tau 1} & 0 \\ 0 & 0 & \sigma^2_{\tau 2} \end{pmatrix}$$

The parameters of the constrained random coefficients matrix $\Sigma_{\eta 1}$ and $\Sigma_{\tau 1}$ are tested using likelihood ratio test following Snijders and Bosker (2011). Using the 50:50 mixture of chi-square distribution with zero and one degrees of freedom, the null that $\sigma^2_{\tau 2} = 0$ could not be rejected, but the null that $\sigma^2_{\eta 2} = 0$ is rejected. This implies that the random component of the quadratic term of nitrogen is significant at the farmer level but not at the field level. Moreover, the null hypothesis that the covariance between the random slope and and random intercept is zero at the field level (i.e. $\upsilon_{01} = 0$) could not be rejected. Therefore, the final random component matrices to be estimated for model four is given by $\hat{\Gamma}_\eta$ and $\hat{\Gamma}_\tau$ as below:

$$\hat{\Gamma}_{\eta 1} = \begin{pmatrix} \sigma^2_{\eta 0} & \rho_{01} & 0 \\ \rho_{10} & \sigma^2_{\eta 1} & 0 \\ 0 & 0 & \sigma^2_{\eta 2} \end{pmatrix} \qquad \hat{\Gamma}_{\tau 1} = \begin{pmatrix} \sigma^2_{\tau 0} & \upsilon_{01} \\ \upsilon_{10} & \sigma^2_{\tau 1} \end{pmatrix}$$

The multilevel models one to four with random components specification $\tilde{\Sigma}_\eta$, and $\hat{\Sigma}_\eta$ and $\hat{\Sigma}_\tau$ , are estimated via REML.

### 4.4.2 Estimation of the fixed components of the regression

The test for the fixed part of the model is implemented using the likelihood ratio test with $\tilde{\Sigma}_\eta$ for models 1 to 3 and $\hat{\Sigma}_\eta$ and $\hat{\Sigma}_\tau$ as the variance-covariance matrix of the random component for model 4. The test for fixed components are based on MLE rather than REML. Under REML specification, the number of fixed component variables are not same across the two models, as it takes into account the degrees of freedom (number of explanatory variables in the fixed part) in the estimation of random parameters. Therefore, the two models in spite of having the same random parameters specification are not comparable under REML, and therefore MLE is used for valid comparison. The final parameter estimates of all the models 1 to 4, with random components specification $\tilde{\Sigma}_\eta$ $\bar{\Sigma}_\eta$ and $\hat{\Sigma}_\eta$ are reported using the REML, as in table 4.4 in the appendix in section 4.6.

### 4.4.3 Discussion

The results of the chapter are discussed in the following three sections. The parameter estimates in table 4.4 in the appendix section 4.6 are used to report the results of the following section. The first section uses results from the multilevel regression to estimate the subjective

MPN, and compares it with the objective benchmark that has been developed in chapter 2. The second section discusses the findings of the multilevel regression results in context of farmers' belief about nitrogen management. The third section reports the skewness of the nitrogen conditional subjective yield distribution that was measured in chapter 2.

### Bias in subjective beliefs of yield nitrogen relationship

A farmer's expected utility maximization is given by (as discussed in chapter 3)

$$\max_{N} \; EU = \int U(\pi(N,W)) \cdot dF(W)$$

where, $\pi(N,W)$ is the farm profit defined as

$$\pi(N,W) = P \cdot f(N,W) - P_N \cdot N$$

$f(N,W)$ is the yield corresponding to $N$, nitrogen and $W$, weather. $P$ and $P_N$ are the corn and nitrogen prices respectively. $U(\pi)$ is the farmer's utility from farm profits, such that $U_\pi > 0$. A farmer is said to be risk-neutral if $U_{\pi\pi} = 0$, and risk-averse if $U_{\pi\pi} < 0$. Under the above framework, optimal nitrogen $N^*$ is determined by

$$P \cdot \int U_\pi \cdot f_N(N^*,W) \cdot dF(W) = P_N \cdot \int U_\pi \cdot dF(W) \tag{4.12}$$

Equation 4.12 describes the choice of optimal nitrogen for an expected utility maximizing farmer. Intuitively it means that the gain in dollar value of expected utility of an additional unit of nitrogen (in terms of the price of corn) is equal to the cost in dollar value of expected utility of an additional unit of nitrogen (in terms of the price of nitrogen). However, for a risk neutral farmer since $U_{\pi\pi} = 0$, equation 4.12 can be written as

$$f_N(N^*,W) = \frac{P_N}{P}$$

The optimal nitrogen for an expected utility maximizing farmer is determined by the following equation:

$$\frac{\int (U_\pi \cdot F_N(N^*,W)) \cdot dF(W)}{\int U_\pi \cdot dF(W)} = \frac{P_N}{P} \tag{4.13}$$

If it is assumed that farmers are expected utility maximizers, then every $k^{th}$ farmer chooses nitrogen $N_{3jk}$ in his $j^{th}$ field according to equation 4.13, such that $N^* = N_{3jk}$. In addition to this, farmers belief about expected yield is elicited at nitrogen levels around $N_{3jk}$. Every $k^{th}$ farmer reports expected yield, $Y_{ijk}$ corresponding to counterfactual nitrogen levels, $N_{ijk}$ for $i = \{1, 2, 4, 5\}$. The two step elicitation of nitrogen and its corresponding expected yield is distinct at each step. In eliciting $N_{3jk}$ from the farmers in the first step, nitrogen is an optimal choice ($N^*$) of farmers, which is an outcome of equation 4.13. Under expected utility maximization, the optimal choice of nitrogen results from the farmer's perception of MPN indexed with the risk preference under weather uncertainty, such that the risk preference weighted MPN is equal to the price ratio of nitrogen and corn (as in equation 4.13). On the other hand, while eliciting expected yield ($Y_{ijk}$) corresponding to counterfactual nitrogen levels ($N_{ijk}$ such that $i = \{1, 2, 4, 5\}$), the nitrogen levels are exogenous[5]. Therefore, it is safe to consider the elicited yield nitrogen response as their subjective belief, since it does not include any choice to be made and hence, it is not unlikely that risk preferences do not play a role here. The elicited values represent the perception of farmers about the nitrogen yield relationship in their field, which are measured around their optimal nitrogen choice.

The parameter estimates from the multilevel regression model 1 in table 4.4, in the appendix, section 4.6 are used to estimate the subjective MPN for each farmer, at their optimal level of nitrogen application, $N_{3jk}$. It is important to note that the survey of subjective expectations elicited the expected yield at different counterfactual levels of nitrogen, which are then used to arrive at the estimates of subjective MPN, denoted by $\frac{\partial Y_{ijk}}{\partial \bar{N}_{ijk}}$. This is likely to minimize errors or any kind of measurement bias that may have arisen in direct elicitation of MPN.

Farmers have some weather expectations under which they have reported their yield expectations. For the purpose of comparison it is assumed that farmers have rational expectations about weather uncertainty. Farmers' perception of weather uncertainty is assumed to be the same as weather distribution over the last fifteen years (from 1999 - 2013) for the Ames farm (as modeled in the previous chapter). The survey respondents are sampled from central Iowa, so

[5]Although they were anchored on the choice of nitrogen made by the farmer, they appeared to the respondent as exogenous

the weather distribution for Ames (referred to as the objective weather distribution) is assumed to be representative of farmers' belief about weather uncertainty. The three weather variables used in modeling the objective yield distribution in the previous chapter are the pre-season precipitation, the growing season precipitation and the total degree days during the growing season (from May to August). The survey of subjective beliefs described in chapter 2, was conducted around the middle of June in 2014. Therefore, the timing of the survey rules out any uncertainty regarding the pre-season precipitation, when farmers reported their subjective yield beliefs. Hence, for the objective weather distribution, the growing season temperature and rainfall are assumed to be the only weather uncertainty associated with the subjective yield expectations. A trivariate log normal distribution is fitted to weather variables in chapter 3, which is used in this chapter to capture the weather uncertainty through the bivariate marginal distribution of the growing season temperature and rainfall. Moreover, it is also assumed that all farmers in the sample have elicited their yield expectations corresponding to the pre-season precipitation values for 2013-2014 weather in Ames.

The weather and nitrogen conditional objective yield and MPN estimates are weighted with the joint probability of the weather variables, calculated from the marginal bivariate log normal weather distribution. The weather weighted objective yield and MPN estimates provide a benchmark of objective yield and MPN, which is conditional on nitrogen and an assumed joint log normal distribution of weather variables, which captures weather uncertainty synonymous to what farmers may have perceived. Under the above assumptions, the subjective expected yield and MPN estimates are compared to the weather weighted objective estimates of yield and MPN.

Figure 4.3 compares the nitrogen conditional subjective yield as reported by the farmers with the weather weighted nitrogen conditional objective yield. The results show that for the CC rotation, most farmers' yield expectations are aligned with the objective nitrogen conditional average yield. But the results for the SC rotation show that the subjective yields are lower compared to the objective yields.

As already mentioned in chapter 2, the response from the subjective expectations survey reveal that most farmers have associated some nitrogen usage to corn yields grown in their field.

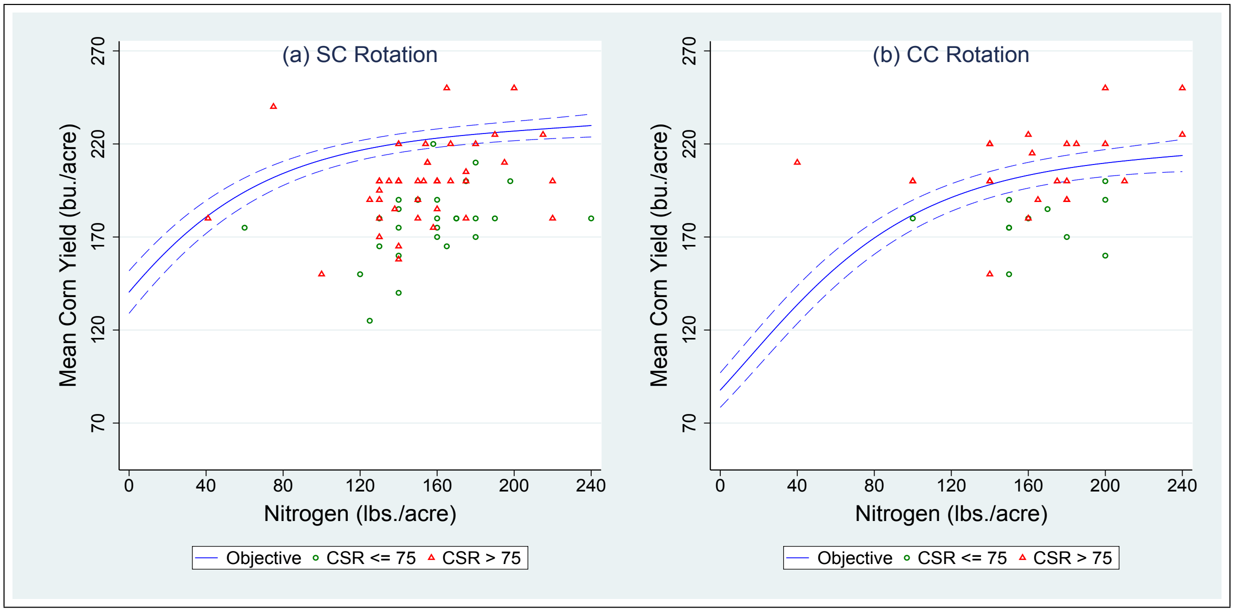



Figure 4.3 Comparison of subjective nitrogen conditional expected yield (denoted by red (triangles) and green (circles)) with the weather probability weighted objective average nitrogen conditional yield corresponding to Ames in central Iowa.

Few farmers in the sample have elicited expected yield values corresponding to corn yields (as response of 100 bu./acre or more is unlikely to be soybean yields), in spite of the fact that they have reported to grow soybeans on their field during the 2014 growing season. Therefore, for all such farmer respondents it is assumed that although, they reported to grow soybean, all of their response corresponding and subsequent to nitrogen usage have been elicited for corn yields (which includes yield expectations and counterfactual yield nitrogen mapping). It has been assumed that they have recorded their response for a hypothetical scenario, if they would have grown corn on their field in the growing season of 2014. Under this assumption, it is not known whether their elicited expectations are conditional on the crop rotation they have originally reported in the beginning of the survey or is it based on an alternative crop rotation they have in their mind while reporting their nitrogen usage and corresponding yield expectations. It is possible that the subjective yield underestimation for SC rotation is a result of this ambiguity, however similar results are also observed for those farmers, who are not a part of this anomaly.

Moreover, farmers were asked to report the N-P-K ratio of their every monthly nitrogen application along with the quantity of monthly nitrogen applied. As already mentioned in chapter 2, while many farmers reported the ratio, some have also reported the amount of nitrogen. Therefore, for few farmers there is no clear indication as to whether the quantity of nitrogen reported is the actual nitrogen content or the amount of fertilizer applied (from which the nitrogen content needs to be estimated based on N-P-K ratio). For the latter case, speculative adjustment has to be made for the nitrogen content based on the information on N-P-K ratio. This could have been a possible reason for the above result. However, the response for farmers for whom there is no such scope for adjustment, also showed similar results as in figure 4.3.

In figure 4.3, subjective yield expectations reported by a farmer corresponding to field with CSR less than or equal to 75 is denoted by green circles and greater than 75 CSR is denoted by red triangles. As it can be seen that farmers who have reported lower yield expectations relative to the objective benchmark (for which the CSR is 76), many of those correspond to lower CSR fields as denoted by red triangles. For the SC rotation, the yield expectations may be lower as farmers may have chosen to follow SC rotation in fields with lower CSR (although the mean CSR for SC and CC rotation is 76 and 77.5 respectively, which is not significantly different statistically).

The average nitrogen application of farmers under SC rotation is 157 lbs./acre, which is significantly lower than the average nitrogen application under CC rotation which stands to be 173 lbs./acre. Besides the difference in the nitrogen application rate, no significant difference in any other field related variables is found. Therefore, there is a possibility that the farmers are pessimistic about expected corn yield under the SC rotation as SC rotation involves use of lesser nitrogen. But if use of lesser nitrogen compared to CC is the cause of underestimation of expected yields, then it is not known why farmers choose less nitrogen in the first place under SC rotation. Therefore, this observation is not conclusive and its validity needs to be considered for further investigation.

Figures 4.4 and 4.5 compares the subjective MPN to the weather weighted objective MPN. Figure 4.4 and figure 4.5 corresponds to estimates of subjective MPN based on model 1 and

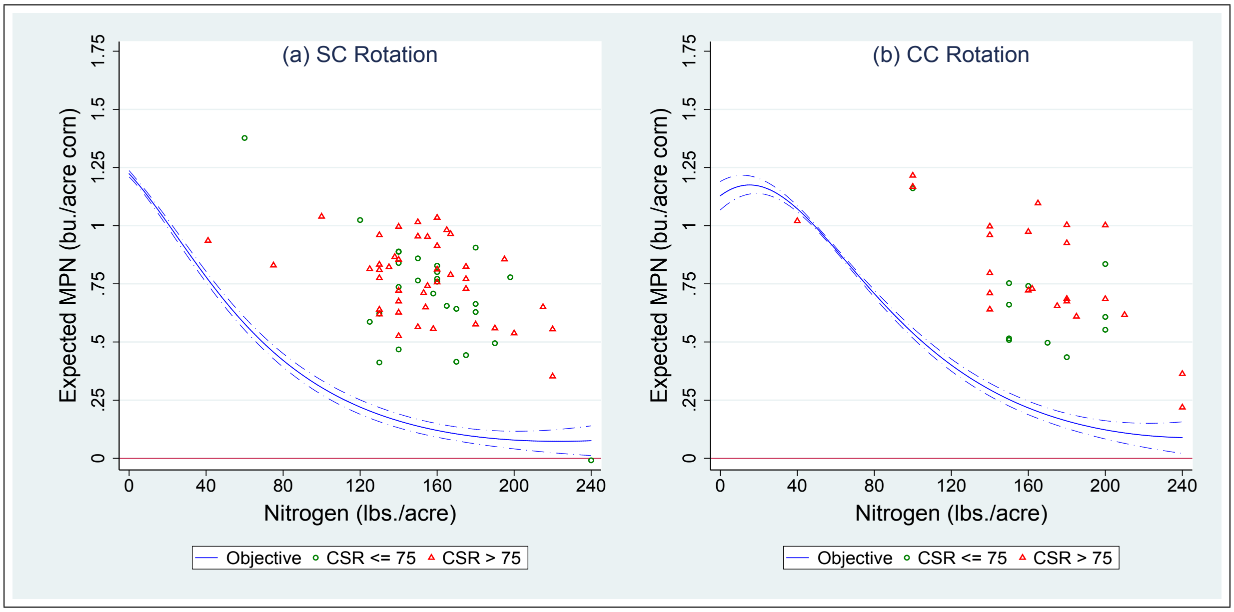



Figure 4.4 Comparison of Subjective MPN (denoted by red (triangles) and the green (circles)) with the weather probability weighted objective MPN corresponding to Ames in central Iowa derived from model 1

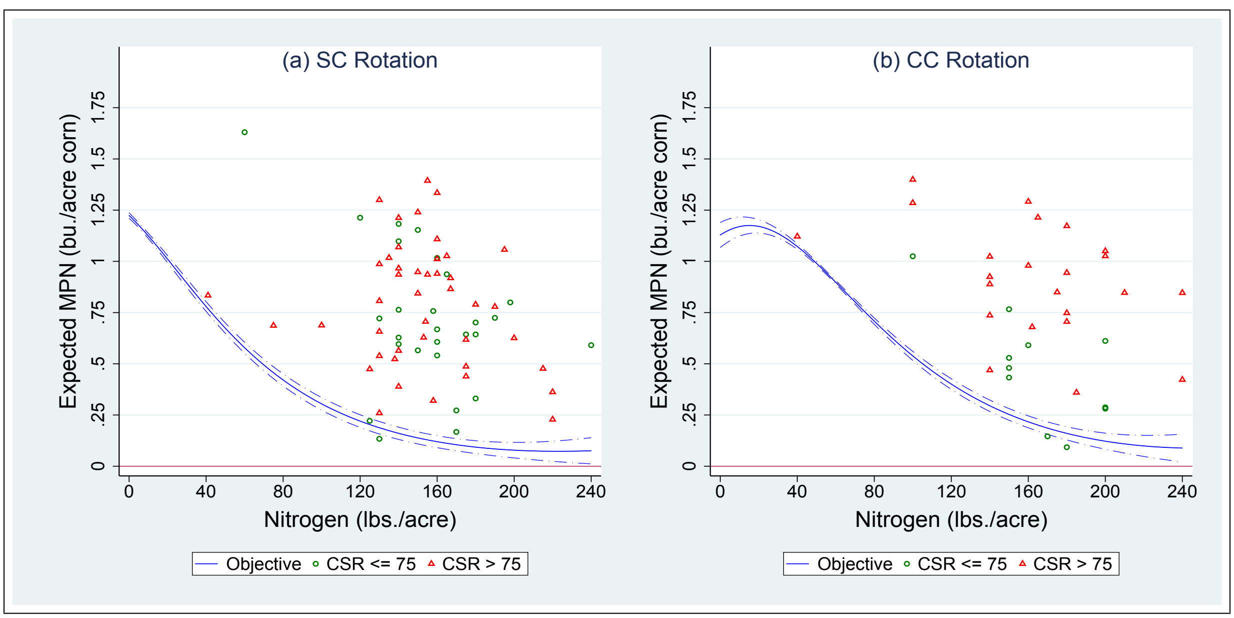



Figure 4.5 Comparison of Subjective MPN (denoted by red (triangles) and the green (circles)) with the weather probability weighted objective MPN corresponding to Ames in central Iowa derived from model 4

model 4 of table 4.4 in appendix 4.6. Both the figures depict that the farmers are upward biased in their beliefs about the MPN across both rotations. The green circles and the red triangles denote subjective MPN corresponding to CSR values less than and greater than equal to 75 respectively. Across both the SC and CC rotation, and all values of CSR, the results seem to indicate the presence of overestimation bias in the farmers' beliefs about MPN. The figures also indicate that the overestimation bias is higher for the more productive fields, which have higher CSR denoted by red triangles.

The survey of subjective expectations was conducted among the farmers in the year 2014. Growing season precipitation during 2014 was relatively higher than the average precipitation (at around $75^{th}$ percentile). The 2014 growing season precipitation was 21.7 inches compared to an average of 19.0 inches. Similarly, the 2014 growing season total degree days was lower at 2370 units than the average of 2445 units. As discussed in chapter 3, both of these weather variable values indicate a relatively favorable weather than the average weather conditions for crop growth. There is a possibility that farmers have taken this into account and have beliefs of higher MPN. However, the best of favorable weather conditions (as described in chapter 3) do not produce objective estimates of MPN, which are even roughly close to the subjective estimates. This indicates that the existence of bias cannot be attributed to a particular weather condition chosen to represent the result.

Another robustness check is conducted to validate the sensitivity of the results against the choice of the multilevel model used. Using the elicited subjective yields for every level of nitrogen application, an individual farmer and field specific OLS regression as described in equation 4.4, is fit to estimate the slope of the nitrogen yield response. The summary of the results of the individual regression are presented in table 4.3.

Figure 4.6 plots the subjective MPN estimates from the individual farmer and field specific OLS linear regression model. It is evident from the figure that the upward bias in the belief about the subjective MPN among farmers exists even while using a farmer and field specific individual linear OLS regression. The results could have been driven by the choice of a particular modeling technique (here multilevel model), but the subjective estimates of the MPN from the individual model confirms the robustness of the result that it is invariant to modeling

Table 4.3 Summary statistics of individual linear regression

| | Mean | Std. dev. | Min. | Max. | Percentiles | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | $5^{th}$ | $25^{th}$ | $50^{th}$ | $75^{th}$ | $95^{th}$ |
| Slope | 0.57 | 0.39 | -0.31e-3 | 2.37 | 0.01 | 0.31 | 0.57 | 0.80 | 1.15 |
| Intercept | 98.66 | 62.72 | -14.58 | 250 | 6.35 | 55.70 | 82.14 | 143.13 | 218.53 |

technique.

All of the above discussion points to the fact that farmers significantly over estimate the MPN around their choice of nitrogen in their fields. As it has been discussed earlier that given the information collected and available, how a farmer arrives at the optimal choice of nitrogen is beyond the scope of this study. This result rather measure the subjective expectations of a farmer around their choice of nitrogen.

**Empirical facts: Subjective beliefs about nitrogen productivity**

The previous subsection presented the evidence of upward bias in farmers' beliefs regarding the MPN. It has been shown that the upward bias in the subjective beliefs of MPN stands out in a field and farmer specific OLS linear regression also. As already discussed, the contribution of multilevel model is to use the correlated information structure within the data while up keeping the advantages of the OLS linear regression. It can identify the contextual effects in the subjective yield regression model. Although, the nitrogen variables, $N_{ijk}$ (except for $i = 3$) are not farmer's choice, and can be treated exogenous, there are no controls for endogeneity of other variables. Therefore, the regression results in table 4.4 are informative about the correlations of the nitrogen variables with other factors across different levels rather than evidence of causal relationship. The effect of $\bar{N}_{ijk}$, which is the amount of nitrogen application centered on the actual application, is the MPN at the average level of nitrogen application. It is interpreted as the within effects of regression.

Farmers who have reported higher than average nitrogen in their field also report higher expected yields, which indicate that across farmers, there is a significant positive associa-

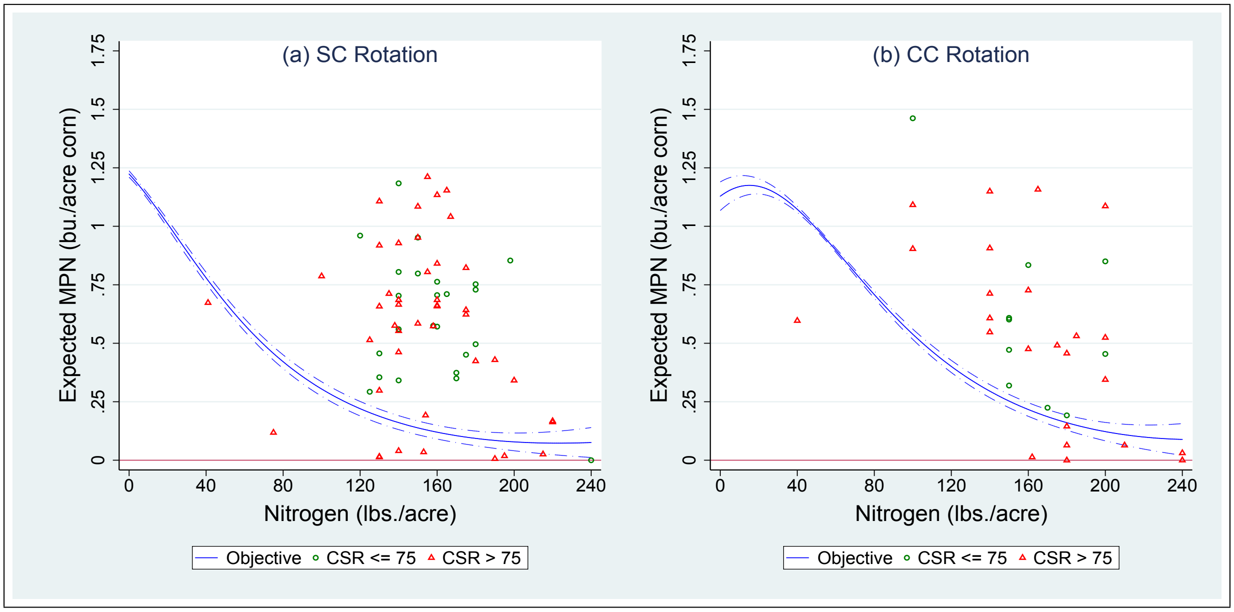



Figure 4.6 Comparison of the weather probability weighted objective MPN with the subjective MPN estimates (denoted by red triangles and the green circles) derived from individual farmer and field specific linear regression

tion between nitrogen and expected yield. Around the average levels of choice of nitrogen, farmer's believe MPN to be significantly positive. Moreover, additional nitrogen corresponding to higher levels of nitrogen application is associated with lower productivity. This indicates that farmers believe MPN to be diminishing in nitrogen. A significant negative coefficient of the $\bar{N}_{ijk} \times Average$ and $\bar{N}_{ijk} \times Under$ is evidence of higher MPN for more productive fields. The cubic term of nitrogen is not statistically significant.

The effect of variables measuring nitrogen productivity during fall and after planting relative to the nitrogen application before planting was not found to be statistically insignificant. This imply that nitrogen productivity do not vary across different timing of application.

Delays in planting and pollination are associated with decline in expected yield across farmers, but a positive coefficient of interaction between nitrogen term with planting date is observed in models 1 to 3. This may be indicative of the fact that higher nitrogen reduces the loss in the expected yield due to delayed planting. It may be argued that farmers with later planting dates have applied more nitrogen. However, the counterfactual nitrogen ($N_{ijk}$) levels are exogenous and therefore, interaction of counterfactual nitrogen levels with the plating date measures the nitrogen productivity around the farmer's choice of optimal nitrogen for a given planting date. Moreover, the interaction between planting date and nitrogen is the contextual effect, which identify the farmers' belief about nitrogen productivity in a given environment (here planting dates). Similarly, the coefficient of interaction term between expected pollination dates and $N_{ijk}$ is negative in model 4, indicating that the nitrogen productivity is lower with delayed expected pollination date (although it is statistically insignificant).

The coefficient of total land farmed and its interaction with nitrogen do not stand out to be statistically significant. However, the percentage of land owned out of total farmed land has a negative and statistically significant coefficient.

Although, the between effects of field size do not stand out to be statistically significant, the positive and statistically significant within effect indicates that for farmers larger fields are associated with higher expected yields.

The negative and statistically significant coefficient of education indicate that farmers educated beyond high school report lower expected yields. Also, farmers who have reported to share their nutrient management decision with a second person are associated with higher expected yields.

**Subjective Yield distribution skewness around nitrogen**

Most studies have advocated that nitrogen increase negative skewness of the crop yield density (Day, 1965; Nelson and Preckel, 1989; Babcock and Hennessy, 1996; Du et al., 2012) and the results from chapter 3 also reconfirms this objective notion.

Chances on a scale from zero to hundred are elicited from farmer respondents about their subjective cumulative probability of yield realization around their expected yield corresponding to their optimal choice of nitrogen. The thresholds used for the left tail are 75% and 90% of the expected yield, and 110% and 125% of expected yield for the right tail. The chances out of 100 (cumulative probability) in the left tail (less than 75% and 90% of expected yield) and the right tail (greater than 125% and 110% of expected yield) are elicited. A crude measure of yield skewness is defined, which is the difference of cumulative probability (chances) assigned in

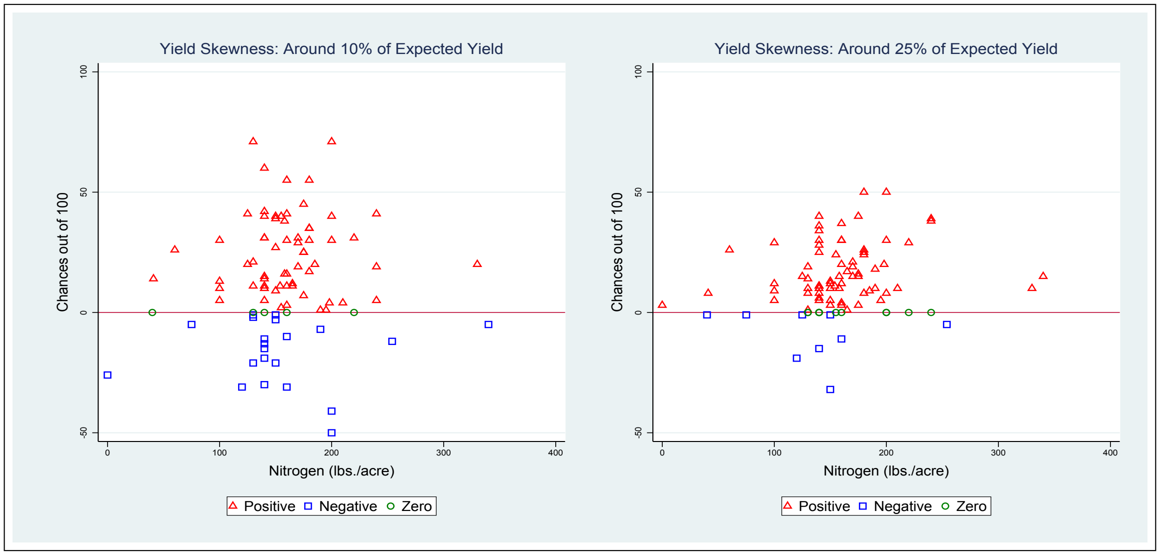



Figure 4.7 Yield skewness measured as the difference of probability mass in the left tail versus the right tail. The left and the right panel measures the yield skewness around 10% and 25% of expected yield respectively, plotted against the amount of nitrogen applied by the farmer. Red triangle, green circle and the blue squares indicate the positive, zero and negative yield skewness.

the left tail over the right tail. The difference in chances are calculated for (less than) 90% and (greater than) 110% (defined as the skewness measure around 10% of expected yield) and, (less than) 75% and (greater than) 125% of the elicited expected yield (called the skewness measure around 25% of expected yield) respectively. Yield distribution is symmetric if the reported differences are zero, indicated by hollow green circles in figure 4.7. Similarly, if the difference is positive (as shown by hollow red triangles) the yield skewness is positive. The blue squares indicate negative yield skewness. The measure of yield skewness is plotted against the optimal amount of nitrogen used (or planned to use) by the farmer respondents in figure 4.7.

In figure 4.8, each bar represents the two yield skewness measure (around 10% and 25% of the expected yield) for a unique field of any farmer. It can be seen that most farmers indicate a positively skewed yield distribution at their chosen level of nitrogen application. Although, figure 4.7 measures the skewness of the yield distribution at the chosen levels of nitrogen application by farmers, it cannot sign the direction of change in the skewness of the yield density followed by a change in nitrogen application amount. It could increase the yield skewness (by reducing positive skewness) as shown in chapter 3 and previous studies or vice-

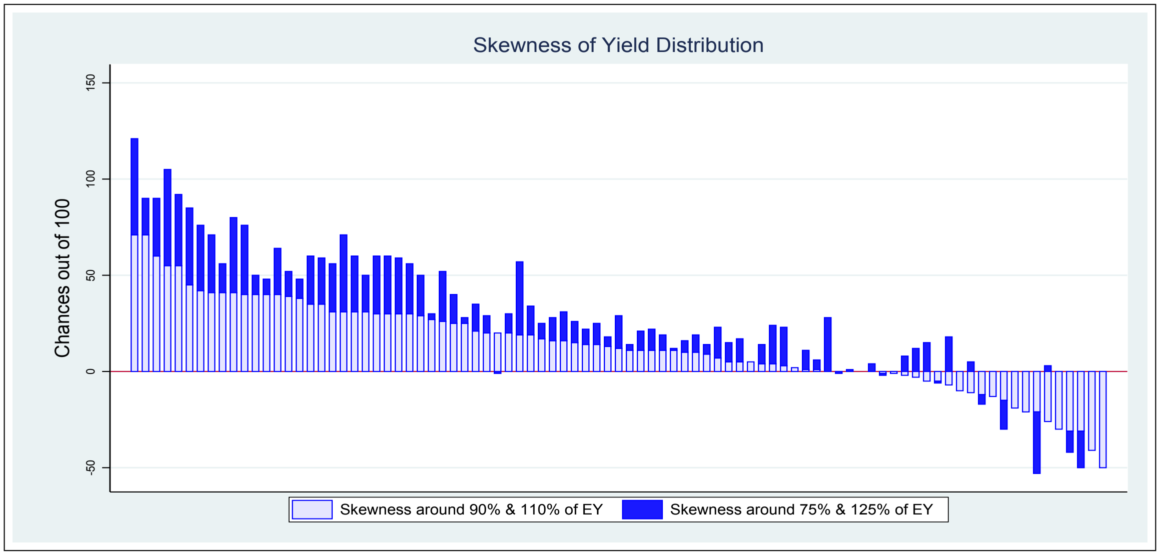



Figure 4.8 Measure of Yield skewness for subjective nitrogen conditional yield distributions for the *Best producing fields*. The light shaded bar indicates the excess weight individual farmers put in the left tail versus the right tail around the 90% and 110% of subjective yield. The dark shaded bar measures the same around 75% and 125% of expected yield.

versa. The observation indicates that at the amount of nitrogen application (around an average of 160 lbs./acre across farmers), the subjective beliefs of farmers are not consistent with the objective model developed in chapter 3 and that of previous studies, which have found negative yield skewness at similar levels of nitrogen application.

## 4.5 Discussion and Conclusion

This chapter uses results from the previous chapters to study farmers' subjective beliefs surrounding nutrient management in their fields in context to the objective model of yield density. The moments of the crop yield distribution of the objective yield density is modeled using data on ex-post realized yields in chapter 3. The survey data described in chapter 2 measured the subjective expectations of farmers around their chosen level of nitrogen.

A multilevel model is used to estimate the nitrogen conditional subjective yield of farmers and the contextual effects associated with them. While comparing the objective nitrogen conditional mean yield with the subjective yield expectations, there is some evidence that farmers underestimate the subjective yields in SC rotation under equivalent levels of nitrogen.

Possible conjectures are made for this observation, but a concrete and definitive answer requires further investigation. Moreover, the estimates of MPN from the multilevel subjective yield model are compared with the MPN estimates from the objective model. This reveals that mostly farmers overestimate the MPN at their chosen levels of nitrogen application. Babcock (1992) mentions that if the mean of the MPN distribution of subjective beliefs are higher than those held objectively, in context to the objectively held belief, the higher subjective mean MPN is evidence of perceived over-application of nitrogen. This research stands testimony to this fact but also acknowledges that nitrogen over-application is rather a relative term and the perception of nutrient over-application is contextual to the underlying beliefs as stated. This chapter refrains from making conclusions about nitrogen over-application. It rather tries to get closer to the reality by presenting the fact that there is substantial divergence between the subjective belief of farmers and the agronomic relationship. The research also finds that subjective yield skewness measured with the cumulative probability of the yield distribution measured in chapter 2 are positive for most farmers compared to the evidence of negative yield skewness at similar or equivalent levels of nitrogen application.

## 4.6 Appendix

Table 4.4 Regression estimates of subjective yield model

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $\bar{N}_k$ | 0.124$^{*}$ | 0.110 | 0.074 | 0.128$^{**}$ |
| | (0.0698) | (0.0694) | (0.0802) | (0.0640) |
| $\bar{N}_{jk}$ | 0.491x10$^{-1**}$ | 0.569x10$^{-1***}$ | 0.357x10$^{-1}$ | 0.448x10$^{-1}$ |
| | (0.0219) | (0.0219) | (0.0233) | (0.0507) |
| $\bar{N}_k \times \bar{N}_{jk}$ | -0.197x10$^{-2***}$ | -0.204x10$^{-2***}$ | -0.159x10$^{-2***}$ | -0.209x10$^{-2***}$ |
| | (0.0341x10$^{-2}$) | (0.0335x10$^{-2}$) | (0.0395x10$^{-2}$) | (0.0794x10$^{-2}$) |
| $\bar{N}_{ijk}$ | 0.743$^{***}$ | 0.712$^{***}$ | 0.855$^{***}$ | 0.905$^{***}$ |
| | (0.0549) | (0.0566) | (0.0929) | (0.0660) |
| $\bar{N}_k \times \bar{N}_{ijk}$ | -0.176x10$^{-2**}$ | -0.186x10$^{-2**}$ | -0.299x10$^{-2***}$ | -0.212x10$^{-2**}$ |
| | (0.0897x10$^{-2}$) | (0.0879x10$^{-2}$) | (0.0967x10$^{-2}$) | (0.0939x10$^{-2}$) |
| $\bar{N}_{jk} \times \bar{N}_{ijk}$ | -0.353x10$^{-2***}$ | -0.349x10$^{-2***}$ | -0.355x10$^{-2**}$ | -0.386x10$^{-2***}$ |
| | (0.0135x10$^{-1}$) | (0.0130x10$^{-1}$) | (0.0146x10$^{-1}$) | (0.0115x10$^{-1}$) |
| $\bar{N}_k \times \bar{N}_{jk} \times \bar{N}_{ijk}$ | 0.407x10$^{-4**}$ | 0.409x10$^{-4**}$ | 0.514x10$^{-4**}$ | 0.461x10$^{-4***}$ |
| | (0.0186x10$^{-3}$) | (0.0179x10$^{-3}$) | (0.0217x10$^{-3}$) | (0.0162x10$^{-3}$) |
| $\bar{N}_{ijk}^2$ | -0.899x10$^{-2***}$ | -0.884x10$^{-2***}$ | -0.996x10$^{-2***}$ | -0.970x10$^{-2***}$ |
| | (0.0734x10$^{-2}$) | (0.0708x10$^{-2}$) | (0.0874x10$^{-2}$) | (0.0737x10$^{-2}$) |
| $\bar{N}_{ijk} \times (Average)_{jk}$ | -0.156$^{***}$ | -0.149$^{***}$ | -0.158$^{***}$ | -0.157$^{***}$ |
| | (0.0517) | (0.0505) | (0.0519) | (0.0510) |
| $\bar{N}_{ijk} \times (Under)_{jk}$ | -0.227$^{***}$ | -0.228$^{***}$ | -0.246$^{***}$ | -0.235$^{***}$ |
| | (0.0537) | (0.0521) | (0.0571) | (0.0522) |

Standard errors in parentheses

$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Table 4.4 (Continued)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $(C\bar{S}R)_k$ | 0.555** | 0.544** | 0.448 | 0.477** |
| | (0.2350) | (0.2360) | (0.3010) | (0.2320) |
| $(Fiel\bar{d}\ Size)_k$ | -0.593x10$^{-1}$ | -0.607x10$^{-1}$ | -0.697x10$^{-1}$ | -0.298x10$^{-1}$ |
| | (0.0541) | (0.0544) | (0.0602) | (0.0515) |
| $(C\bar{S}R)_{jk}$ | 1.106*** | 0.968*** | 1.120*** | 1.140*** |
| | (0.0634) | (0.0659) | (0.0697) | (0.1440) |
| $(Fiel\bar{d}\ Size)_{jk}$ | 0.944x10$^{-1}$*** | 0.099*** | 0.101*** | 0.101*** |
| | (0.0167) | (0.0163) | (0.0169) | (0.0381) |
| $(Planting)_{jk}$ | -4.166** | -5.494*** | -5.216** | |
| | (2.0040) | (1.9770) | (2.1030) | |
| $(Pollination)_{jk}$ | -8.283*** | -7.640*** | -7.380*** | -4.476** |
| | (1.6920) | (1.6970) | (1.8250) | (1.9100) |
| $(Pollination)_{jk}$ | -8.283*** | -7.640*** | -7.380*** | -4.476** |
| | (1.6920) | (1.6970) | (1.8250) | (1.9100) |
| $(Education)_k$ | -9.045** | -8.056* | -9.677* | |
| | (4.435) | (4.454) | (5.163) | |
| $(Shared\ decision)_k$ | 20.31*** | 16.52*** | 20.13*** | 17.27*** |
| | (4.735) | (5.125) | (5.543) | (4.530) |
| $(Land\ \bar{f}armed)_k$ | 0.129 | 0.128 | 0.180 | 0.188 |
| | (0.2000) | (0.1990) | (0.2150) | (0.1790) |
| $(\%\ Land\ owned)_k$ | -0.125** | -0.125** | -0.149** | -0.107** |
| | (0.0567) | (0.0570) | (0.0661) | (0.0537) |
| $(Planting)_{jk} \times \bar{N}_{ijk}$ | 0.086** | 0.795x10$^{-1}$** | 0.721x10$^{-1}$* | |
| | (0.0410) | (0.0404) | (0.0393) | |
| $(Land\ \bar{f}armed)_k \times \bar{N}_{ijk}$ | -0.423x10$^{-2}$ | -0.402x10$^{-2}$ | | -0.476x10$^{-2}$* |
| | (0.0286x10$^{-1}$) | (0.0278x10$^{-1}$) | | (0.0261x10$^{-1}$) |
| $(Average)_{jk}$ | | -7.647*** | | |
| | | (1.7430) | | |

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4 (Continued)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $(CSC\ rotation)_{jk}$ | | -7.776*** | -5.810** | |
| | | (1.9540) | (2.2870) | |
| $(Shared\ decision)_k \times \bar{N}_{ijk}$ | | 0.104* | | |
| | | (0.0620) | | |
| $\bar{N}^3_{ijk}$ | | | 0.469x10$^{-4}$ | |
| | | | (0.0295x10$^{-3}$) | |
| $(Nitrogen\ Test)_{jk}$ | | | 14.650*** | |
| | | | (4.0070) | |
| $(Center\ Prob)_{jk}$ | | | 0.163** | |
| | | | (0.0716) | |
| $(Center\ Prob)_{jk} \times \bar{N}_{ijk}$ | | | -0.346x10$^{-2}$** | |
| | | | (0.0139x10$^{-1}$) | |
| $(Pollination)_{jk} \times \bar{N}_{ijk}$ | | | | -0.321x10$^{-1}$ |
| | | | | (0.0255) |
| $Constant$ | 206.600*** | 212.900*** | 196.800*** | 191.700*** |
| | (5.2180) | (5.2900) | (6.9930) | (4.7290) |
| **Random effects:** | | | | |
| **Farmer level** | | | | |
| $Var(\bar{N}_{ijk})$ | 0.0288x10$^{-1}$*** | 0.281x10$^{-1}$*** | 0.256x10$^{-1}$*** | 0.360x10$^{-1}$*** |
| | (0.1670) | (0.1670) | (0.1920) | (0.1660) |
| $Var(Intercept)$ | 347.537*** | 342.044*** | 419.646*** | 197.759*** |
| | (0.107) | (0.104) | (0.112) | (0.158) |
| $Cov(Intercept,\ \bar{N}_{ijk})$ | -1.875*** | -1.696*** | -2.062*** | -1.356** |
| | (0.2510) | (0.2350) | (0.2730) | (0.2800) |
| $Var(\bar{N}^2_{ijk})$ | | | | 0.118x10$^{-4}$*** |
| | | | | (0.2540) |
| **Random effects:** | | | | |
| **Field level** | | | | |
| $Var(\bar{N}_{ijk})$ | | | | 0.013*** |
| | | | | (0.3090) |
| $Constant$ | | | | 146.129*** |
| | | | | (0.1260) |

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4 (Continued)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Residual Variance** | | | | |
| $Var(\epsilon_{ijk})$ | 148.836*** | 137.335*** | 139.633*** | 71.916*** |
| | (0.0351) | (0.0351) | (0.0376) | (0.0427) |
| $N$ | 540 | 540 | 480 | 540 |
| $BIC$ | 4721.0 | 4699.5 | 4237.6 | 4580.9 |

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

# CHAPTER 5. CONCLUSION

Nutrient decision making is at the center of agricultural decision making. The literature in agronomy and agricultural economics has linked nitrogen decisions to other aspects of agriculture which includes field productivity and crop rotation, risk management, insurance decision, farm practices, etc.

This dissertation investigate subjective beliefs of farmers, which are likely to be an essential component of the nutrient decision making. This dissertation provides new results for understanding nutrient management decisions under uncertainty. It takes a first step toward a complete understanding of the nutrient management choices of farmers. However, how a farmer arrives at the decision of how much nitrogen to apply on his field is beyond the scope of this dissertation. Rather, the contribution of this dissertation is to measure farmer's perception about the nitrogen yield relationship and uncertainty in general, which is likely to motivate the choice of optimal nitrogen. Moreover, the measured subjective beliefs gather farmer's perception of the yield nitrogen relationship, which is difficult to be measured from farmer's observed choices or their participation in experimental studies (as their perception are likely to get contaminated by their risk preference).

## 5.1 Discussion of Results

Chapter 2 of the dissertation describes the survey and data collected. The innovation in the survey design is the use of recent advancements in behavioral economics and psychology to measure the subjective beliefs of farmers surrounding the nitrogen management decision in their fields. The unique approach to measure the probability beliefs and expectations do not have precedence in such a context or elsewhere in agricultural economics. The survey responses show that the methodology adopted for the measurement of subjective expectations has been successful in gathering information about the nutrient management decisions. Although, there has been a few glitches in the survey, and some ambiguity in the responses, the ambiguity in the response itself has been informative about certain aspects of the farmer's behavior and perception which stands in contrast to conventionally assumed beliefs. This has provided with potential scope for improvement in the methods and broader applications of the measurement of farmers' subjective expectations.

Chapter 3 of the dissertation has followed the steps of previous research studies that have modeled objective yield distributions. This model has been referred to as the objective model, as estimates from this model are used as an objective benchmark for comparison of of subjective beliefs. The yield distribution for a sample of research farms at ISU is modeled using a generalized linear model with beta distribution and the polynomial terms of the covariates, which provides a flexible structure for the estimation of the yield distribution. Although, the use of a beta generalized linear model is not a precedent in the modeling of yield distribution, the estimated moments of the yield distribution are conditional on both nitrogen and weather variables, which is not done elsewhere. Moreover, it also highlights the economic implications of the weather effect on the MPN and the choice of nitrogen. The modeling technique runs parallel to Just and Pope production function (Just and Pope, 1978, 1979) by estimating a separate sub model for yield variability that provides it the flexibility to model yield variance. Results from this section on yield variance do not meet the conventional wisdom that nitrogen is a risk increasing input as outlined in Just and Pope (1978, 1979).

Chapter 4 builds a subjective model of yield expectations using farmers elicited survey expectation data from chapter 2. A subjective yield expectations model is fit using a multilevel model and the marginal productivity of nitrogen (MPN) is estimated. The comparison of the subjective and the objective estimates of the MPN portrays a significant divergence between the two. This is evident of the fact that farmers perceive MPN to be much higher than the agronomic modeled rates.

As already mentioned that given the information available, it is not possible to identify the exact nitrogen decision-making process of the farmer. Given the choice of nitrogen for a farmer, what is measured are the subjective beliefs of farmers around their chosen nitrogen management practice. In order to identify the nutrient decision-making process, farmer's risk preferences should be measured and combined with the measured subjective expectations. However, possible implications of these measured subjective expectations under alternative risk preferences will be discussed. The nitrogen corn price ratio is assumed to be 0.1 (given by \$0.50 : \$5.00). Under the framework of expected profit maximization, the optimal nitrogen is given by equating the expected MPN to the nitrogen corn price ratio. For most farmers, the subjective estimates of MPN at their chosen level of nitrogen are much higher than the nitrogen corn price ratio of 0.1. Under expected utility maximization framework, the optimal choice of nitrogen is such that the gain in dollar value of expected utility of an additional unit of nitrogen (in terms of the price of corn) is equal to the cost in dollar value of expected utility of an additional unit of nitrogen (in terms of the price of nitrogen). Moreover, for a risk neutral farmer, the optimal nitrogen is same as the optimal nitrogen in the expected profit maximizing framework. Therefore, farmer's nitrogen decision-making seems inconsistent with the expected utility maximization under risk neutrality. For a risk averse farmer, the utility can be denoted by a concave function. With positive marginal utility and risk aversion, the subjective MPN estimates for most farmers cannot rationalize the choice of nitrogen even for extremely risk averse farmer for the given 0.1 nitrogen corn price ratio (or lower). All of the above discussion has been in context of the assumption that farmers operate under the expected utility framework to choose optimal level of nitrogen. However, the expected MPN estimates suggest that the farmer's nitrogen choice is inconsistent with their beliefs under expected utility maximization. Their choice of nitrogen may be driven by alternative preference structure, which could be reference dependence or preference over higher moments of the profit (or yield or wealth). Given the information at hand, one can only conjecture but nothing concrete can be concluded about the choice of nitrogen made.

Subjective beliefs of farmers regarding the yield density show that at the chosen levels of nitrogen, the measure of yield skewness is positive for most farmers. As summarized by previous studies, skewness of yield distribution is positive at low levels of nitrogen, which eventually becomes negative as level of nitrogen application increase. The change in the yield skewness from positive to negative in the objective model indicates that increased nitrogen application creates a more favorable environment for the farmer. This is not evident in subjective beliefs of farmers. The measured subjective beliefs do not allow to comment on the change in yield skewness following higher nitrogen application. But for any given level of nitrogen application, for most farmers, positive value of the measure of yield skewness indicates that farmers do not believe that at the given levels of nitrogen the risk environment is favorable (denoted by higher cumulative probabilities of yield greater than average yield relative to cumulative probabilities of yield lower than average yield). Therefore, in context to the objective yield model, it is not known, whether farmers believe that increasing nitrogen puts them in a more favorable risk environment or not. It may be possible that as nitrogen application increase, the subjective positive yield skewness decrease or increase, but at similar comparable levels of nitrogen there is evidence that farmers do not believe the risk environment to be favorable as suggested by the objective model.

## 5.2 Future research

The implications of research findings are discussed in the previous section. It can be seen that as mentioned previously, the research study aims to contribute to the measurement of subjective beliefs of farmers around nitrogen management practices.

The comparison of subjective beliefs with the objective benchmark has brought attention to the divergence between the subjective beliefs and the objective benchmark, which states the value in measuring subjective expectations. The discordance of subjective beliefs with the objective benchmark has raised questions in this context, which are research questions that will be of interest for future studies.

The inconsistency of subjective expectations of MPN with expected utility framework suggests future research in identifying the alternative decision-making framework, which is used by farmers.

The contrast revealed in the asymmetry of the crop distribution, is of significant importance to production risk. It becomes all the more relevant with the demand for crop insurance in the picture.

Studies have found that corn producers in the Midwest adopt Best Management Practices (BMP) at low rates. Measuring of subjective expectations (and perceived risk) associated with the BMP adoption can bring in information about low observed rates of BMP adoption.

# REFERENCES


Allcott, H. (2013). The welfare effects of misperceived product costs: Data and calibrations from the automobile market. *American Economic Journal: Economic Policy*, 5(3):30–66.

Alpert, M. and Raiffa, H. (1982). A progress report on the training of probability assessors. In Kahneman, D., Slovic, P., and Tversky, A., editors, *Judgment Under Uncertainty: Heuristics and Biases*, pages 294–305. Cambridge University Press.

Antle, J. M. (1983). Testing the stochastic structure of production: a flexible moment-based approach. *Journal of Business & Economic Statistics*, 1(3):192–201.

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

Attanasio, O. P. (2009). Expectations and perceptions in developing countries: their measurement and their use. *The American Economic Review*, 99(2):87–92.

Atwood, J., Shaik, S., and Watts, M. (2003). Are crop yields normally distributed? a reexamination. *American Journal of Agricultural Economics*, 85(4):888–901.

Babcock, B. A. (1992). The effects of uncertainty on optimal nitrogen applications. *Review of Agricultural Economics*, 14(2):271–280.

Babcock, B. A. and Blackmer, A. M. (1992). The value of reducing temporal input nonuniformities. *Journal of Agricultural and Resource Economics*, pages 335–347.

Babcock, B. A. and Hennessy, D. A. (1996). Input demand under yield and revenue insurance. *American journal of agricultural economics*, 78(2):416–427.

Berck, P. and Helfand, G. (1990). Reconciling the von liebig and differentiable crop production functions. *American Journal of Agricultural Economics*, 72(4):985–996.

Blass, A. A., Lach, S., and Manski, C. F. (2010). Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability*. *International Economic Review*, 51(2):421–440.

Boyer, C. N., Brorsen, B. W., and Tumusiime, E. (2015). Modeling skewness with the linear stochastic plateau model to determine optimal nitrogen rates. *Agricultural Economics*, 46(1):1–10.

Boyer, C. N., Larson, J. A., Roberts, R. K., McClure, A. T., Tyler, D. D., and Zhou, V. (2013). Stochastic corn yield response functions to nitrogen for corn after corn, corn after cotton, and corn after soybeans. *Journal of Agricultural and Applied Economics*, 45(4):669.

Campbell, D., Carruth, A., Dickerson, A., and Green, F. (2007). Job insecurity and wages. *The Economic Journal*, 117(518):544–566.

Carlson, G. A. (1970). A decision theoretic approach to crop disease prediction and control. *American Journal of Agricultural Economics*, 52(2):216–223.

Carlson, R. E. (1990). Heat stress, plant-available soil moisture, and corn yields in iowa: A short-and long-term view. *Journal of Production Agriculture*, 3(3):293–297.

Cerrato, M. and Blackmer, A. (1990). Comparison of models for describing; corn yield response to nitrogen fertilizer. *Agronomy Journal*, 82(1):138–143.

Chang, H.-H. and Mishra, A. K. (2012). Chemical usage in production agriculture: Do crop insurance and off-farm work play a part? *Journal of environmental management*, 105:76–82.

Chen, C.-C., McCarl, B. A., and Schimmelpfennig, D. E. (2004). Yield variability as influenced by climate: a statistical investigation. *Climatic Change*, 66(1-2):239–261.

Chesley, G. (1978). Subjective probability elicitation techniques: A performance comparison. *Journal of Accounting Research*, pages 225–241.

Claassen, R. and Just, R. E. (2011). Heterogeneity and distributional form of farm-level yields. *American journal of agricultural economics*, 93(1):144–160.

Clop-Gallart, M. and Juárez-Rubio, F. (2007). Elicitation of subjective crop yield pdf [probability density functions]. *Spanish Journal of Agricultural Research*, 5(1):16–24.

Day, R. H. (1965). Probability distributions of field crop yields. *Journal of Farm Economics*, 47(3):713–741.

Delavande, A. (2008). Pill, patch, or shot? subjective expectations and birth control choice*. *International Economic Review*, 49(3):999–1042.

Delavande, A. (2014). Probabilistic expectations in developing countries. *Annu. Rev. Econ.*, 6(1):1–20.

Delavande, A., Giné, X., and McKenzie, D. (2011a). Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design? *Journal of Applied Econometrics*, 26(3):479–497.

Delavande, A., Giné, X., and McKenzie, D. (2011b). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics*, 94(2):151–163.

Delavande, A. and Kohler, H.-P. (2009). Subjective expectations in the context of hiv/aids in malawi. *Demographic research*, 20:817.

Delavande, A. and Kohler, H.-P. (2012). The impact of hiv testing on subjective expectations and risky behavior in malawi. *Demography*, 49(3):1011–1036.

Diez, R. (2002). A glossary for multilevel analysis. *Journal of epidemiology and community health*, 56(8):588.

Dominitz, J. (2001). Estimation of income expectations models using expectations and realization data. *Journal of Econometrics*, 102(2):165–195.

Dominitz, J. and Manski, C. (1996). Eliciting student expectations of the returns to schooling. *Journal of Human Resources*, 31(1):1–26.

Dominitz, J. and Manski, C. F. (1997a). Perceptions of economic insecurity: Evidence from the survey of economic expectations. *Public Opinion Quarterly*, 61(2):261–287.

Dominitz, J. and Manski, C. F. (1997b). Using expectations data to study subjective income expectations. *Journal of the American statistical Association*, 92(439):855–867.

Dominitz, J. and Manski, C. F. (2004). How should we measure consumer confidence? *The Journal of Economic Perspectives*, 18(2):51–66.

Du, X., Hennessy, D. A., and Cindy, L. Y. (2012). Testing days conjecture that more nitrogen decreases crop yield skewness. *American Journal of Agricultural Economics*, 94(1):225–237.

Du, X., Yu, C. L., Hennessy, D. A., and Miao, R. (2015). Geography of crop yield skewness. *Agricultural Economics*, 46(4):463–473.

Enders, C. K. and Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, 12(2):121.

Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.

Fuller, W. A. (1965). Stochastic fertilizer production functions for continuous corn. *Journal of Farm Economics*, 47(1):105–119.

Gallagher, P. (1986). Us corn yield capacity and probability: estimation and forecasting with nonsymmetric disturbances. *North Central Journal of Agricultural Economics*, pages 109–122.

Gouret, F. and Hollard, G. (2011). When kahneman meets manski: Using dual systems of reasoning to interpret subjective expectations of equity returns. *Journal of Applied Econometrics*, 26(3):371–392.

Grisley, W. and Kellogg, E. D. (1983). Farmers' subjective probabilities in northern thailand: an elicitation analysis. *American Journal of Agricultural Economics*, 65(1):74–82.

Hampton, J., Moore, P., and Thomas, H. (1973). Subjective probability and its measurement. *Journal of the Royal Statistical Society. Series A (General)*, pages 21–42.

Haran, U., Moore, D. A., and Morewedge, C. K. (2010). A simple remedy for overprecision in judgment. *Judgment and Decision Making*, 5(7):467.

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

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

Hennessy, D. A. (2011). Modeling stochastic crop yield expectations with a limiting beta distribution. *Journal of Agricultural and Resource Economics*, pages 177–191.

Hofmann, D. A. and Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 24(5):623–641.

Horowitz, J. K. and Lichtenberg, E. (1993). Insurance, moral hazard, and chemical use in agriculture. *American journal of agricultural economics*, 75(4):926–935.

Huber, G. P. (1974). Methods for quantifying subjective probabilities and multi-attribute utilities. *Decision Sciences*, 5(3):430–458.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2).

Just, R. E. and Pope, R. D. (1978). Stochastic specification of production functions and economic implications. *Journal of econometrics*, 7(1):67–86.

Just, R. E. and Pope, R. D. (1979). Production function estimation and related risk considerations. *American Journal of Agricultural Economics*, 61(2):276–284.

Just, R. E. and Weninger, Q. (1999). Are crop yields normally distributed? *American Journal of Agricultural Economics*, 81(2):287–304.

Kaylen, M. S. and Koroma, S. S. (1991). Trend, weather variables, and the distribution of us corn yields. *Review of agricultural economics*, 13(2):249–258.

Ker, A. P. and Coble, K. (2003). Modeling conditional yield densities. *American Journal of Agricultural Economics*, 85(2):291–304.

Lawley, C., Lichtenberg, E., and Parker, D. (2009). Biases in nutrient management planning. *Land Economics*, 85(1):186–200.

Lindsey, J. K. (1997). *Applying generalized linear models*. Springer Science & Business Media.

Ludke, R. L., Stauss, F. F., and Gustafson, D. H. (1977). Comparison of five methods for estimating subjective probability distributions. *Organizational Behavior and Human Performance*, 19(1):162–179.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5):1329–1376.

Manski, C. F. and Straub, J. D. (2000). Worker perceptions of job insecurity in the mid-1990s: Evidence from the survey of economic expectations. *Journal of Human Resources*, pages 447–479.

McCullagh, P. and Nelder, J. A. (1989). *Generalized linear models*, volume 37. CRC press.

McDonald, J. B. and Xu, Y. J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics*, 66(1):133–152.

Menapace, L., Colson, G., and Raffaelli, R. (2013). Risk aversion, subjective beliefs, and farmer risk management strategies. *American Journal of Agricultural Economics*, 95(2):384–389.

Miranda, M. J. and Fackler, P. L. (2004). *Applied computational economics and finance*. MIT press.

Mishra, A. K., Nimon, R. W., and El-Osta, H. S. (2005). Is moral hazard good for the environment? revenue insurance and chemical input use. *Journal of environmental management*, 74(1):11–20.

Moss, C. B. and Shonkwiler, J. (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.

Norris, P. E., Kramer, R. A., et al. (1990). The elicitation of subjective probabilities with applications in agricultural economics. *Review of Marketing and Agricultural Economics*, 58(2-3):127–147.

Norwood, B., Roberts, M. C., and Lusk, J. L. (2004). Ranking crop yield models using out-of-sample likelihood functions. *American Journal of Agricultural Economics*, 86(4):1032–1043.

Paolino, P. (2001). Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Analysis*, 9(4):325–346.

Park, W. I. and Sinclair, T. R. (1993). Consequences of climate and crop yield limits on the distribution of corn yields. *Review of Agricultural Economics*, 15(3):483–493.

Paulson, N. D. and Babcock, B. A. (2010). Readdressing the fertilizer problem. *Journal of Agricultural and Resource Economics*, pages 368–384.

Pingali, P. L. and Carlson, G. A. (1985). Human capital, adjustments in subjective probabilities, and the demand for pest controls. *American Journal of Agricultural Economics*, 67(4):853–861.

Preckel, P. V., Shively, G. E., Baker, T. G., Chu, M.-C., and Burrell, J. E. (2000). Contract incentives and excessive nitrogen use in agriculture. *Journal of Agricultural and Resource Economics*, pages 468–484.

Rajsic, P. and Weersink, A. (2008). Do farmers waste fertilizer? a comparison of ex post optimal nitrogen rates and ex ante recommendations by model, site and year. *Agricultural Systems*, 97(1):56–67.

Ramirez, O. A. (1997). Estimation and use of a multivariate parametric model for simulating heteroskedastic, correlated, nonnormal random variables: the case of corn belt corn, soybean, and wheat yields. *American Journal of Agricultural Economics*, 79(1):191–205.

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

Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*, volume 1. Sage.

Roberts, M. J., Schlenker, W., and Eyer, J. (2012). Agronomic weather measures in econometric models of crop yield with implications for climate change. *American Journal of Agricultural Economics*, page aas047.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. a definition. *Journal of Economic theory*, 2(3):225–243.

Sawyer, J., Nafziger, E., Randall, G., Bundy, L., Rehm, G., and Joern, B. (2015). Concepts and rationale for regional nitrogen rate guidelines for corn. *Iowa State UniversityUniversity Extension PM*.

Sawyer, J. E. (2016). Nitrogen use in iowa corn production.

Schaefer, R. E. and Borcherding, K. (1973). The assessment of subjective probability distributions: A training experiment. *Acta Psychologica*, 37(2):117–129.

Schlenker, W. and Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. *Applied Economic Perspectives and Policy*, 28(3):391–398.

Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37):15594–15598.

Seaver, D. A., Von Winterfeldt, D., and Edwards, W. (1978). Eliciting subjective probability distributions on continuous variables. *Organizational Behavior and Human Performance*, 21(3):379–391.

Sheriff, G. (2005). Efficient waste? why farmers over-apply nutrients and the implications for policy design. *Applied Economic Perspectives and Policy*, 27(4):542–557.

Sherrick, B., A. Lanoue, C., Woodard, J., D. Schnitkey, G., and D. Paulson, N. (2014). Crop yield distributions: fit, efficiency, and performance. *Agricultural Finance Review*, 74(3):348–363.

Sherrick, B. J. (2002). The accuracy of producers' probability beliefs: Evidence and implications for insurance valuation. *Journal of Agricultural and Resource Economics*, pages 77–93.

Sherrick, B. J., Sonka, S. T., Lamb, P. J., and Mazzocco, M. A. (2000). Decision-maker expectations and the value of climate prediction information: conceptual considerations and preliminary evidence. *Meteorological Applications*, 7(4):377–386.

Sherrick, B. J., Zanini, F. C., Schnitkey, G. D., and Irwin, S. H. (2004). Crop insurance valuation under alternative yield distributions. *American Journal of Agricultural Economics*, 86(2):406–419.

Smith, V. H. and Goodwin, B. K. (1996). Crop insurance, moral hazard, and agricultural chemical use. *American Journal of Agricultural Economics*, 78(2):428–438.

Smithson, M. and Merkle, E. C. (2013). *Generalized linear models for categorical and continuous limited dependent variables*. CRC Press.

Smithson, M. and Verkuilen, J. (2006). A better lemon squeezer? maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*, 11(1):54.

Snijders, T. A. and Bosker, R. J. (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE.

SriRamaratnam, S., Bessler, D. A., Rister, M. E., Matocha, J. E., and Novak, J. (1987). Fertilization under uncertainty: an analysis based on producer yield expectations. *American Journal of Agricultural Economics*, 69(2):349–357.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *The Review of economic studies*, 81(1):426–472.

Tack, J., Harri, A., and Coble, K. (2012). More than mean effects: modeling the effect of climate on the higher order moments of crop yields. *American Journal of Agricultural Economics*, 94(5):1037–1054.

Taylor, C. R. (1984). A flexible method for empirically estimating probability functions. *Western Journal of Agricultural Economics*, pages 66–76.

Tembo, G., Brorsen, B. W., Epplin, F. M., and Tostão, E. (2008). Crop input response functions with stochastic plateaus. *American Journal of Agricultural Economics*, 90(2):424–434.

Thompson, L. M. (1986). Climatic change, weather variability, and corn production. *Agronomy Journal*, 78(4):649–653.

Thompson, L. M. (1988). Effects of changes in climate and weather variability on the yields of corn and soybeans. *Journal of Production Agriculture*, 1(1):20–27.

Tumusiime, E., Wade Brorsen, B., Mosali, J., Johnson, J., Locke, J., and Biermacher, J. T. (2011). Determining optimal levels of nitrogen fertilizer using random parameter models. *Journal of Agricultural and Applied Economics*, 43(4):541.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131.

Winkler, R. L. (1967). The assessment of prior distributions in bayesian analysis. *Journal of the American Statistical association*, 62(319):776–800.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, 48(3):545–595.

Zhu, Y., Goodwin, B. K., and Ghosh, S. K. (2011). Modeling yield risk under technological change: Dynamic yield distributions and the us crop insurance program. *Journal of Agricultural and Resource Economics*, pages 192–210.