

2009

Economic benefits of the National Cooperative Soil Survey Program

Archana Pradhan
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Pradhan, Archana, "Economic benefits of the National Cooperative Soil Survey Program" (2009). *Graduate Theses, Dissertations, and Problem Reports*. 4514.
<https://researchrepository.wvu.edu/etd/4514>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Economic Benefits of the National Cooperative Soil Survey Program

Archana Pradhan

**Dissertation submitted to the
Davis College of Agriculture, Forestry, and Consumer Sciences
at West Virginia University
in partial fulfillment of the requirements
for the degree of**

**Doctor of Philosophy
in
Natural Resource Economics**

**Jerald J. Fletcher, Ph. D., Chair
Tim Phipps, Ph. D.
Alan Collins, Ph. D.
Michael P. Strager, Ph. D.
Stratford Douglas, Ph. D.**

**Agricultural and Resource Economics Program
Division of Resource Management**

Morgantown, West Virginia

2009

Key words: Soil Survey Information, Economic Benefits, Crop Yield

ABSTRACT

Economic Benefits of the National Cooperative Soil Survey Program

Archana Pradhan

The National Cooperative Soil Survey (NCSS) is the primary source of information on the soil resources in the U.S. The information provided by the NCSS program has played a significant and important role in diverse fields. This study estimates the net benefits of the information provided by the soil survey program to the production of selected crops. Benefit estimates are based on relative productivity gains related to the provision of soil information at the county level. The estimated value of increased crop yields less estimated soil survey production costs provides a lower-bound estimate of the total economic benefits of the NCSS.

The structure of the NCSS program provides a spatial-temporal pattern to the development of county level soil information that can be interpreted as a natural experiment where the outcomes provide a means of estimating a partial benefit of the value of soil survey information in agriculture production. Benefit-cost ratios are utilized to evaluate the effectiveness of the NCSS program.

A benefit-cost analysis of the NCSS for the corn and soybeans production regions based on a 7% discount rate gave a benefit-cost ratio of 7:1 for the correlation date scenario and 5:1 for the publication date scenario. This suggests that even the lower bound estimate of benefits based on productivity increases for just two crops, corn and soybeans, outweighs the cost of the entire soil survey program for the study region.

The results from the benefit-cost analysis suggest that the NCSS program is economically viable in areas of the country considered. This is a promising result given the incomplete nature of the currently available data. In summary, this research seeks to compute a lower-bound estimate of the economic benefits of the NCSS for major crops and thus contribute to the documentation of the value of information provided by the NCSS.

ACKNOWLEDGEMENTS

I would like to thank, first and foremost, my committee chair Dr. Jerald J. Fletcher for allowing me to work on this project and for his critical review, intellectual perspective, constructive criticism, and guidance. I am very much indebted for his guidance and advice in every step of difficult moments for the completion of this research. I would also like to thank Dr. Timothy Phipps, Dr. Alan Collins, Dr. Michael Strager, and Dr. Stratford Douglas to be a part of this dissertation committee and for their critical comments, review of the manuscript, and helpful insights.

I would like to acknowledge the funding provided for this research by the U. S. Department of Agriculture (USDA)-NRCS through the NGDC. I would like to thank Jon Hempel, Sharon Waltman, and Henry Ferguson from the NGDC-NRCS for providing necessary data and information during the research period. I would also like to thank Jeffrey Andresen from Michigan State Climatologist's Office for providing weather data and Paul Benedict from the NCSS for providing NCSS cost.

My personal thanks goes to Kelly Street, Melanie Jimmie, Ellen Hartley-Smith, and Lisa Lewis for their administrative assistance and willingness to help at any time. I would also like to thank Marc Mulkey for his expert editing. I am also grateful to the Department of Resource Management, West Virginia University for the education and experience I have received during my Ph.D. Program. Lastly, I would like to recognize all my friends and professors at the Division of Resource Management for their significant role in this academic experience and my family members for their encouragement and moral support.

DEDICATION

I would like to dedicate this dissertation to my mother Mrs. Shashi K. Pradhan and my father Mr. Shanta K. Pradhan, who instilled in me the value of higher education, desire to learn and the importance of hard work. Without their continuous teaching and encouragement, this endeavor would not have been initiated.

I am especially grateful to my husband Mr. Sandeep Pyakurel for his constant support and inspiration in every step of my dissertation. Most importantly, I also dedicate this work to my daughter Shrina Pyakurel for her patience and love throughout this period of study.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iii
Dedication	iv
Table of Contents	v
List of Figures	vii
List of Tables	viii
Chapter 1	1
Introduction.....	1
1.1. Problem Statement.....	1
1.2. Research Objectives.....	3
1.3. Methodology	3
1.4. Organization of the Study	4
Chapter 2.....	5
Literature Review.....	5
2.1. Introduction.....	5
2.2. Previous Studies on the Value of Information.....	5
2.2.1. Case Studies of Value of Information.....	7
2.2.2. Conclusion	11
2.3. Previous Studies on Cost and Benefits of Soil Survey Information	11
2.4. Previous Crop Yield Studies.....	15
2.5. Summary	21
Chapter 3.....	23
Theoretical Background – Valuation of Soils Information.....	23
3.1. Introduction.....	23
3.2. Soil Information	23
3.2.1. Development of Second Order Soil Survey.....	28
3.3. Value of Information	32
3.3.1. Theoretical Framework:.....	34
3.4. Value of Soil Information	39
3.4.1. Benefits of Soil Survey Information in Agriculture Production.....	40
3.4.2. Economic Analysis of Availability of Soil Information	44
3.4.3. Soil Information as a Public Good.....	46
3.4.4. Nonmarket Valuation Approaches to Valuing Soil Survey Information.....	48
3.5. Natural Experiment.....	51

Chapter - 4.....	53
Empirical Models and Description of Data	53
4.1. Introduction.....	53
4.2. Statistical Model of Crop Production	54
4.2.1. Fixed Effects Method.....	60
4.2.2. Data Description	63
4.2.3. Crop Yield Model	69
4.3. Benefit-Cost Analysis	79
4.3.1. Methodological Approach	80
4.3.2. Choice of Discount Rate	81
4.3.3. Evaluation Technique	82
4.3.4. Benefit Estimation	84
4.3.5. Cost Estimation of Soil Information.....	85
Chapter 5.....	87
Empirical Results and Analysis	87
5.1. Empirical Estimation from Crop Models.....	87
5.1.1. Corn Yield Model	88
5.1.2. Soybean Yield Model	90
5.1.3. Wheat Yield Model.....	92
5.1.4. Cotton Yield Model	97
5.1.5. Validity of Fixed Effects Model	101
5.2. Results from Benefit-cost and Sensitivity Analysis	104
Chapter 6.....	111
Summary and Conclusions	111
6.1. Summary and Implications of the Research	111
6.2. Policy Implications	117
6.3. Limitations and Future Research	118
6.3.1. Limitations	118
6.3.2. Future Research	121
References.....	122
Appendix A.....	129
Appendix B	132
Appendix C.....	133
Appendix D.....	137

LIST OF FIGURES

Figure 2.1: Relation between Cost and Values with Purity	14
Figure 2.2: Crop Yield Trends for Major Crops (1935 = 1).....	16
Figure 2.3: Crop Yield Trends for Major Crops (Unit Bushels /Acre).....	17
Figure 3.1: Users of Soil Survey Information.....	24
Figure 3.2: Soil Geography Hierarchy Diagram.....	26
Figure 3.3: Status Map of the U.S. Second-Order Soil Survey	27
Figure 3.4: Flow Diagram Showing the Users of Soil Survey Information	28
Figure 3.5: Simplified Decision Tree for Farmer with Land ‘A’ and Land ‘B’ (Assuming Soil Survey is Perfect Information).....	42
Figure 3.6: A Farm Lying on Four Soil Types (A-D), with Different Response Curves for a Particular Crop	43
Figure 3.7: Welfare Analysis in Market Equilibrium Framework.....	45
Figure 4.1: Correlation Date for Major Corn Producing Counties in the U.S.....	55
Figure 4.2: Publication Date for Major Corn Producing Counties in the U.S.....	56
Figure 4.3: Average Yield for Major Corn Producing Counties in the U.S. in 2007	56
Figure 4.4: Scatter Plot of County Size vs. Correlation Date of Soil Survey	57
Figure 4.5: Scatter Plot of Crop Acreage vs. Correlation Date of Soil Survey	57
Figure 4.6: Scatter Plot of Average Yield and Dissemination of Soil Survey (Using Correlation Date)	58
Figure 4.7: Selected Corn Producing Counties.....	77
Figure 4.8: Selected Soybean Producing Counties.....	77
Figure 4.9: Selected Winter Wheat Producing Counties	78
Figure 4.10: Selected Spring Wheat Producing Counties.....	78
Figure 4.11: Selected Cotton Producing Counties.....	79
Figure 5.1: Net Benefit Graph for Corn and Soybeans.....	108
Figure 5.2: Total Benefit and Cost Graph for Corn and Soybeans.....	108
Figure 5.3: State Level Benefit-Cost Ratio for Corn and Soybeans Region (Correlation Date Scenario).....	110

LIST OF TABLES

Table 4.1: Correlation Matrix - Provision of Soil Survey Information, Crop Acreage and Crop Yield	55
Table 4.2: USDA-NASS Yield Data Status on Winter Wheat	63
Table 4.3: USDA-NASS Yield Data Status on Cotton.....	64
Table 4.4: Number of Soil Surveys Correlated for the Corn Belt States	65
Table 4.5: Definition and Data Source for Variables Used in Crop Model.....	71
Table 4.6: Top Ten U.S. Corn Producing States in 2007	72
Table 4.7: Top ten U.S. Soybeans Producing States in 2007	73
Table 4.8: Top ten U.S. Winter Wheat Producing States in 2007	73
Table 4.9: Top ten U.S. Spring Wheat Producing States in 2007.....	74
Table 4.10: Top ten U.S. Upland Cotton Producing States in 2007	74
Table 4.11: Descriptive Statistics - Corn Yield Model.....	75
Table 4.12: Descriptive Statistics – Soybeans Yield Model.....	76
Table 4.13: Descriptive Statistics – Wheat and Cotton Yield Model	76
Table 4.14: Basic Assumptions for Cash Flow Analysis.....	81
Table 4.15: Estimates of State Level Federal Budget Allocation of the NCSS Program.....	86
Table 5.1: Number of observations included in each model	87
Table 5.2: Corn Yield Model Results (Publication Date).....	89
Table 5.3: Corn Yield Model Results (Correlation Date).....	90
Table 5.4: Soybeans Yield Model Results (Publication Date)	91
Table 5.5: Soybeans Yield Model Results (Correlation Date)	92
Table 5.6: Winter Wheat Yield Model Results (Publication Date)	94
Table 5.7: Winter Wheat Yield Model Results (Correlation Date)	95
Table 5.8: Spring Wheat Yield Model Results (Publication Date).....	96
Table 5.9: Spring Wheat Yield Model Results (Correlation Date).....	97
Table 5.10: Upland Cotton Yield Model Results (Publication Date)	98
Table 5.11: Upland Cotton Yield Model Results (Correlation Date)	99
Table 5.12: Summary of Increase in Crop Yield Attributable to the Provision of Soil Information	100
Table 5.13: Estimates from Group-Mean Centered Method	102

Table 5.14: Sum of Benefits and Cost for Corn and Soybeans	104
Table 5.15: Sum of Benefits and Cost for Cotton	105
Table 5.16: NPV and Benefit-Cost for Corn andSoybeans	105
Table 5.17: NPV and Benefit-Cost for Cotton.....	106
Table 5.18: State Level Benefit-Cost Ratio for Corn and Soybeans Region (Correlation Date Scenario).....	108
Table 5.19: State Level Benefit-Cost Ratio for Corn and Soybeans Region (Publication Date Scenario).....	108
Table 5.20: State Level Benefit-Cost Ratio for Cotton Region (Correlation Date Scenario).....	110
Table 5.21: State Level Benefit-Cost Ratio for Cotton Region (Publication Date Scenario).....	110

CHAPTER 1

INTRODUCTION

1.1. PROBLEM STATEMENT

The National Cooperative Soil Survey (NCSS) program is a cooperative effort of federal, state, and county agencies. The NCSS is the primary source for collecting and providing soil data for the U.S. This program carries out its activities on national, regional, and state levels under the leadership and coordination of the U.S. Department of Agriculture (USDA), Natural Resources Conservation Service (NRCS). The soil survey program was formally initiated in 1899 (Smith, 1998). Early soil surveys focused on the capabilities of land for agriculture production. Soil surveys conducted after World War II are considered modern soil surveys and have been completed for most of the private land in the U.S. under the NCSS program.

The primary goal of the soil survey program is to assist society and individuals to understand the suitability and limitations of the soil resources for intended uses (Ditzler, Engel, and Ahrens, 2003). The information contained in soil survey inventories has played a significant role in increasing the productivity of the agricultural system, in reducing environmental damage, and other numerous sectors as well (USDA-NRCS, 2009). However, there is little work done on the value of or the benefits derived from the information provided by soil surveys. Since the comprehensive soil survey program is nearing completion, it is appropriate to review the contributions of the program. Estimates of the values of historical and current benefits as well as potential future benefits from

additional investments would facilitate program management and guide policy decisions that will determine future investments.

Soil information has been helpful to farmers to manage and better understand crop growth. For example, Klingebiel (1966) reported that farmers in Hall County, Nebraska had extra income because of the availability of soil information, which helped them to improve water management and reclaim saline land. He also found in one case that the income of a farmer in Fayette County, TN was increased by more than \$5500 in single year as a result of management changes related to soil information. Thus agriculture is one sector that benefits directly from the availability of soil survey information.

This study analyzes the effects of the availability of modern soil survey information on major crop yield trends for major crop producing counties in the U.S. The initial focus is on corn for counties in the Corn Belt. The estimates obtained from this study serve as a partial measure of the benefits due to the increased availability of soils information provided by the NCSS program. Value of the increased crop yield attributable to the availability of soil information provides an estimate of the economic benefits generated by the use of the soil information in crop production. The benefit estimates are comprehensive reflecting changes on both the intensive and extensive margins that result from increased availability of soil information. This research expands the existing literature by providing statistically reliable estimates of the benefits from soil survey information derived from data on aggregate agricultural production.

1.2. RESEARCH OBJECTIVES

The primary objective of this study is to develop a partial assessment of the benefits that accrue to the NCSS through the production of selected agricultural crops.

Specific study objectives are as follows:

- Develop econometric methods to estimate the increase in corn production due to availability of soil information in major corn producing counties, which includes:
 - identifying the data and appropriate statistical techniques, and
 - using the estimates to develop measures of value.
- Test the methods developed for corn in applications to soybeans, cotton, and wheat.
- Understand the production factors that differ by crop and incorporate those drivers in the crop model.
- Aggregate the benefits for the four major crops included in the analysis.
- Conduct an ex-post, partial benefit-cost analysis of the NCSS program which provides a lower bound estimate for social gains from this program.

1.3. METHODOLOGY

This study focuses primarily on estimating the economic benefits of the NCSS program. In so doing, it attempts to empirically test relationships between crop yield and the availability of soil information provided by the NCSS. To capture the changes in crops yield attributable to the provision of soil survey information, a crop yield model is developed. The crop yield model is based on the prior knowledge that crop yield is

dependent on weather, soil productivity, and the level of technology. It is hypothesized that the availability of soils information provided by the NCSS is also an important factor in aggregate yield.

The statistical analysis provides estimates of the change in yield attributable to the information provided by the soil survey program. The monetary value of the change in crop yield based on historical crop prices gives an initial estimate of the economic benefits of the soil survey information. The benefit estimates obtained from the crop yield model provide information to conduct a benefit-cost analysis of the soil survey program.

This study adds to the existing literature on the valuation of soil survey information and to the economic literature on value of information.

1.4. ORGANIZATION OF THE STUDY

This study includes five additional chapters. Chapter 2 provides a literature review of the value of information, benefit-cost analysis of soil survey information, and crop yield studies. Chapter 3 provides the theoretical background on valuation of soil information. Chapter 4 discusses specification of the empirical model on benefit estimation, benefit-cost analysis, and of the nature and sources of data. Chapter 5 provides analysis of results from crop models for benefit estimation and benefit-cost analysis. Finally, Chapter 6 provides a summary, conclusions, policy implications of the results, and limitations of this study.

CHAPTER 2

LITERATURE REVIEW

2.1. INTRODUCTION

This chapter provides a review of the literature pertinent to this research. The first section reviews previous studies on the value of information. The second section discusses previous studies on estimating the value of provision of soil information, and benefit-cost analysis of soil survey information. It also discusses the relationships among costs, benefits, the map scale of the soil survey and the effort level required to produce the soil survey. The last section provides a review of selected analyses of crop yield trends. The models and variables related to increasing crop productivity are discussed for corn, wheat, cotton, and soybeans.

2.2. PREVIOUS STUDIES ON THE VALUE OF INFORMATION

There have been few empirical studies in the field of information economics. Stiglitz (2000) discussed the contributions of information economics to the field of economics. He stated that there remains a significant area for further research in the economics of information and noted that in addition to the monetary value in information, information economics has changed the way economists think.

Most of the limited published papers on information economics are theoretical. Hirsleifer (1973) provided a review of the literature published before the 1970s. Hilton (1981) discussed the main determinants of information value. He synthesized the possible results from these determinants. Chavas and Pope (1984) discussed the measurement and

economic valuation of information. A simplified model was developed to demonstrate how better information could enhance the decision-making process. They stated that the supply of information depends on information cost and features of the query process while demand and the value of information depend on how the economic decision could be improved by the information.

Relatively few papers have dealt with the use of applied response research in determining the value of information. A number of issues make measuring the value of information complicated. Measurement is one issue that makes the economic analysis of information difficult to assess (Chavas and Pope, 1984). Because of its subjective nature, it is difficult to quantify and value information directly. One reason is that information possesses many of the characteristics of public goods. Information is not usually traded in markets like marketable goods. Consumption of information is non-rival and usually non-excludable. Thus quantifying the value of information is difficult because it involves decisions the individual would have made without information and the consequences of those decisions.

While most articles related to the value of information provide theoretical views rather than application, Repo (1989) discussed some of the approaches to estimate the value of information. Perrin (1973) discussed the concept of value of information and applied it to estimate the value of soil test information to corn response research in Brazil.

Roe and Antonovitz (1985) introduced the terms 'willingness to pay' and 'willingness to accept' as a money metric value to estimate the economic value of information. They developed an analytical model for a restricted class of utility functions and applied it by fitting the model to time series data from the U.S. fed cattle industry

assuming a risk-averse scenario. Preckel et al. (1987) extended the model developed by Roe and Antonovitz (1985) to demonstrate an approach to compute the money metric value of information for microeconomic production choices under risk. They used the model to value production information in agricultural production and suggested that the value of information thus obtained could be used to make benefit-cost analysis for the provision of public information.

Repo (1989) stated that few case studies have been able to document empirical evidence of the value of information. Most of the papers dealing with the value of information rely on sensitivity analysis and are based on probabilistic and economic assumptions. Some of them are based on experimental evidence and some are hypothetical. The following section discusses some case studies that have estimated the economic value of information.

2.2.1. Case Studies of Value of Information

2.2.1.1. Weather Information

The most common application of valuing information is in estimating the economic value of weather forecasting. Weather forecast information helps decision makers mitigate adverse consequences that arise from weather effects.

An early attempt was the work of Lave in 1963. Lave investigated the value of weather forecasts to a California raisin farmer whose profit depends on the amount of precipitation in certain stages of grape production. The weather information helps improve the farmer's choice of the optimal picking time. Lave considered climatological probabilities as a prior distribution and employed a decision tree analysis approach to

determine the value of weather information. He found the value of informed decisions to be \$314.65 per acre.

Adams et al. (1995) assessed the economic value of improved forecasts of El Nino weather phenomenon to agriculture production in the southeast U.S. The aggregated economic value on society's payoff was estimated by measuring the total producer and consumer surplus based on meteorological, agronomic, and economic effects. The estimated value for improved forecast information was \$96 million compared to perfect information valued at \$144.5 million. Costello et al. (1998) assessed the value of El Nino-Southern Oscillation (ENSO) in the management of salmon. The ENSO weather forecast provides information on interannual variability in the global climate system. The improved weather information helps producers avoid adverse climatic situations and optimize harvest levels and operations of fish hatcheries. A composite bio-economic model was developed for a Coho salmon fishery to derive the value of information from improved El Nino weather forecasts. The study found that a perfect El Nino forecast provides an annual welfare gain of approximately \$1 billion while imperfect information gains would be smaller.

Likewise, Solow et al. (1998) assessed the economic value of long-range weather prediction by measuring the increase in social welfare resulting from incorporating the ENSO prediction in economic decisions. They used an integrated model that combined meteorology, plant science, and economics in a Bayesian decision approach. They estimated an annual economic value of perfect ENSO predictions to U.S. agriculture to be \$323 million.

Babcock's (1990) study of the value of weather information in market equilibrium contradicts the findings of others. Despite the fact that information is generally considered to increase commodity supply and producer's welfare, improved weather information does not necessarily imply an increase in commodity supply or farmers welfare. Under an inelastic demand, improved weather information might signal farmers to reduce production. Assuming farmers are risk-neutral, rational, and competitive, farmers would not use the information if it did not improve their payoffs. In a competitive market an individual farmer's supply decision does not affect output price. In this situation, improved weather information could have lower value.

2.2.1.2. Soil Test Information

Application of the value of information to valuing soil test information comes after the valuation of weather information. A number of researchers have attempted to assess value of soil test information and discuss related issues in agriculture production (Mitchell, 2003; Babcock, Carriquiry, and Stern, 1966). These studies have shown that soil test information can be valuable to producers. Most of the studies on valuing soil test information are limited to valuing nitrogen soil tests that help the producer decide how much fertilizer should be applied. Past studies indicate that incorporating soil test information could reduce nitrogen fertilizer applications and thus reduce production costs.

Perrin (1976) used two types of response models, the linear response and plateau model (LRP) and the generalized quadratic model, to estimate the value of soil test and soil classification information. He used data from 61 experiments on the fertilizer response of corn for 1967-1969. While the LRP model is not common for economists, it is commonly adopted by agronomists. Results from the quadratic response model implied a

value of soil test information of \$6.16 per hectare while the LRP model implied a value of \$30.92 per hectare. However, the model provided negative estimates for the soil classification information. That is, the payoff to the producer of using soil information was lower compared to the alternative without the use of soil classification information.

Adams et al. (1983) conducted an ex-post assessment of the value of soil test information on nitrogen application in sugar beet production. The data were collected over four years of experiments conducted under irrigation at the Northern Plains Research Center. A two-stage procedure was used to determine the value of soil test information. The first stage involved estimating response functions using physical data on sugar beets for total nitrogen. Price and cost data were then combined with the physical information and economically optimal nitrogen levels were incorporated into the response function for each year to predict output and to estimate ex-post returns. Using soil test information increased producer's returns up to \$62 per acre compared to decisions made without the use of soil information.

Swinton and King (1994) developed an integrated multidisciplinary approach to estimate the value of information. Their research on estimating the value of weed scouting information for management employed a bio-economic, weed management, stochastic simulation model. The bio-economic model was multi-temporal to incorporate all pre and post weed control treatment in the analysis. The value was measured assuming various levels of scouting information for corn and soybeans in southwest Minnesota. They found that incorporating information on weed management could be significant in improving the expected payoffs compared to the payoffs from a fixed decision rule (i.e., without weed information).

2.2.2. Conclusion

Past studies on the value of information have shown that information can be valuable to producers. Most papers discussed the basic contributions of information in increasing payoffs and showed that information has value and helps improve decisions.

Most of the published studies are based on a Bayesian decision analysis framework with a prior probability scenario, and developed using either hypothetical scenarios or experimental evidence. A wide variety of approaches were used ranging from sophisticated econometric analysis to simulation models. Most studies estimated the ex-ante valuation of information. Only a few of the case studies attempted to analyze the ex-post value of information. The available ex-post valuation analyses are based on hypothetical scenarios or experiments that are comparatively small in size in terms of the area and time period involved. The results of such analyses cannot be generalized to a regional or national level.

2.3. PREVIOUS STUDIES ON COST AND BENEFITS OF SOIL SURVEY INFORMATION

There has been little research conducted on the benefits derived from the provision of soil information provided by NCSS. Klingebiel (1966) claimed that investment in soil surveys would be able to pay for the program within a year. He estimated benefit-cost ratios for soil survey investments based on the intensity of land use: a) low intensity (predominantly range and woodland), b) medium density (mixed agriculture and about half cropland), and c) high intensity (rapidly growing metropolitan areas). Benefits of soil information increase with increasing land use intensity. He developed benefit-cost estimates of was 46:1 for low intensity areas, 61:1 for medium intensity areas, and 123:1

for high intensity areas. Estimates of the benefits were determined based on case histories and the records of soil survey users, assuming that most people in the surveyed area would use soil information. According to Klingebiel, there were some cases with much higher benefits that would exceed the estimated ratio. For example, the town of Cohasset in Massachusetts had saved more than \$250,000 by using soil maps while selecting sewage disposal system sites.

Bie and Beckett (1971) found the following relationships between the map scale and the cost of soil survey and effort, respectively. The relationship is expressed by:

$$\log E = 7.41 + 1.57 \log S$$

$$\log C = 8.16 + 1.4 \log S$$

where, E is the effort in man-days per km² required in field,

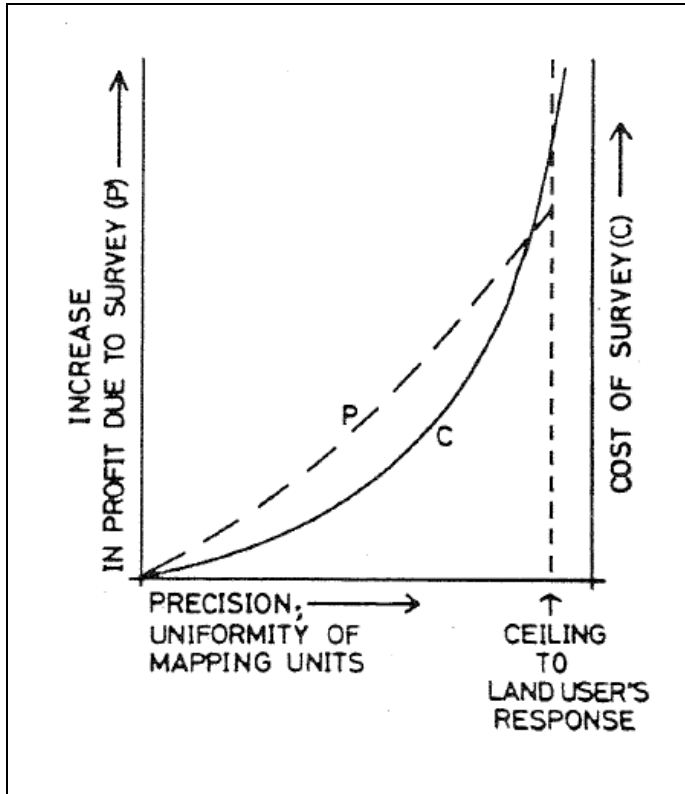
C is the cost of soil survey in U.S. dollar in 1960, and

S is the scale of map.

This equation suggests that the cost of the soil survey is positively related to the scale of the map. Doubling the scale of the map would increase the cost of the soil survey 2.6 fold.

Beckett and Burrough (1971) also suggested that the cost and benefits of the soil surveys rise sharply with increasing quality. Quality here refers to the scale of map where a larger scale gives more detailed information. Bie and Ulph (1972) showed that the value of soil survey information depends on the quality of the maps developed and differences in payoffs among alternative management practices. Their study, based on varieties of peaches, illustrated that gross returns increase as the quality of the information of each mapping unit increases.

Bie et al. (1973) demonstrated a simple algorithm for calculating the possible benefits from further soil survey efforts needed to produce a higher quality map. Western (1978) defined the soil survey value as the ratio of survey quality to survey cost. He emphasized that the term 'survey quality' could have different meanings for the users than the producers of soil surveys, and that it is extremely difficult to quantify the benefits of soil survey information. Beckett (1981) stated that the cost of soil information increases with additional precision and detail. He also noted that it is a complex process to assess the benefits of a soil survey. The general form of the relationship between the cost of a soil survey and the benefits derived increases with the uniformity of mapping units following the law of diminishing returns (Figure 2.1). The uniformity that a map could provide increases with survey cost, and the benefit from soil survey increases with the uniformity of the map's unit. However, the degree of increase is not the same as illustrated in Figure 2.1.



(Source: Beckett, 1981)

Figure 2.1: Relation between Cost and Values with Purity

Dent and Young (1981) stated that U.S. and Australian studies demonstrated benefit-cost ratios from 40:1 to 50:1. They illustrated a simplified example of a methodological approach to estimate the economic benefit of a soil survey: comparing the profitability from different management systems on each of a number of mapping units. This example could be applied to any number of soil units and management units. Highest total profitability could be achieved if each soil unit is positioned on the management type that is most suited for it, and to achieve this requires a soil map and knowledge of the best management for each soil unit.

More recently, Giasson et al. (2000) used the example explained and presented by Dent and Young (1981) to illustrate the analysis for valuing soil information. Decision trees, Bayes' Theorem, and map quality evaluation procedures were used to evaluate the

economic value of soil surveys from three different scenarios, which differed in the level of information concerning soil changes. The three scenarios considered were: (i) site-specific soil information is unavailable, (ii) perfect site-specific soil information is available (not realistic), and (iii) imperfect site-specific soil information is available. The cost of the soil survey was estimated following the relationship given by Bie and Beckett (1971) and converting the soil survey cost to January 2000 U. S. dollars. They derived an estimated economic value of \$17.14/hectare each year from a hypothetical soil survey, which exceeded the estimated soil survey cost of \$2.09.

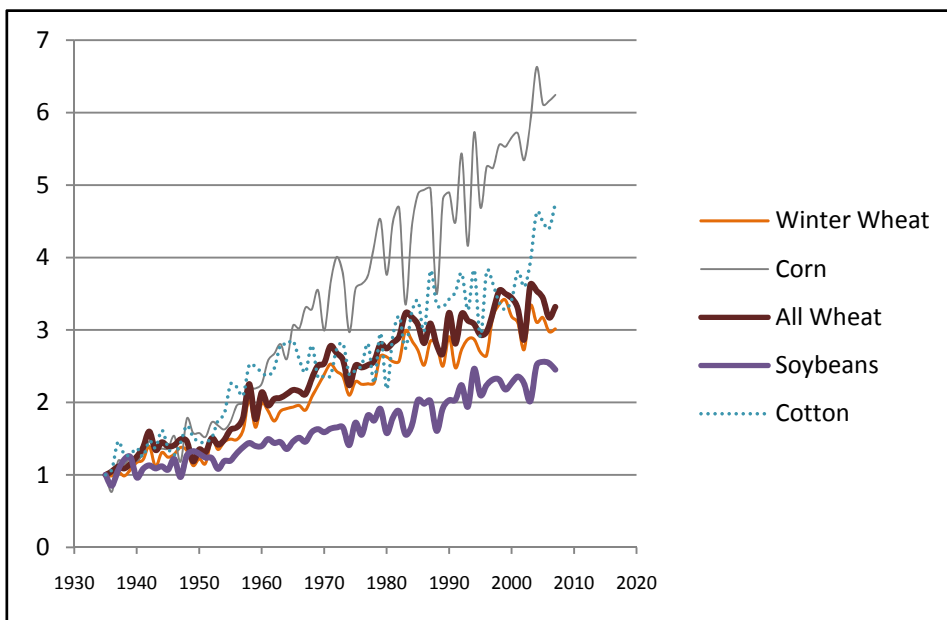
The methods for assessing the value of soil survey information were also based on hypothetical scenario or experimental evidence, which is similar to estimating value of information in the previous sections.

2.4. PREVIOUS CROP YIELD STUDIES

Since this study seeks any evidence of the effect of soil information on increasing crop productivity, it is essential to review previous literature that explains the possible factors that affect crop productivity.

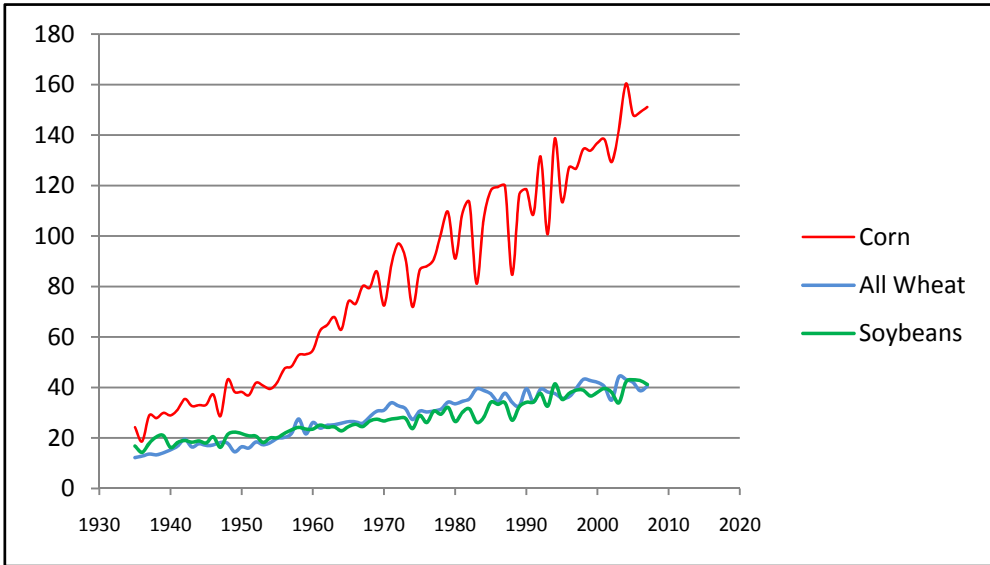
Crop yields have increased dramatically in the U.S. Corn yield in the U.S. averaged 24 bushels per acre in 1935 but had increased by about six fold to 151 bushels per acre in 2007 (USDA-NASS, 2008). Cotton yield increased from 185 pounds per acre in 1935 to almost 900 pounds per acre in 2007. Soybean and wheat yield increases have been less dramatic, increasing by 2-3 fold. Soybean yield increased from 16.8 bushels per acre in 1935 to 41.2 bushels per acre in 2007, and wheat yield increased from 12.2 bushels per acre in 1935 to 40.5 bushels per acre in 2007 (See Figure 2.2 and 2.3).

A number of researchers have attempted to estimate the influence of weather and technology on the increase in crop productivity over a long period using a variety of techniques. Generally, two types of approaches have been employed to assess the impact of weather on crop yields: crop growth simulation models and statistical models. Most studies have used a model with a single-equation framework (Huff and Neill, 1982; Offutt, Garcia and Pinar, 1987; Kaufmann and Snell, 1997).



(Source: USDA-NASS, 2008)

Figure 2.2: Crop Yield Trends for Major Crops (1935 = 1)



(Source: USDA-NASS, 2008)

Figure 2.3: Crop Yield Trends for Major Crops (Unit Bushels /Acre)

There has been extensive research and documentation for corn yield in relation to technology and weather. However, there has been little research on soybean yield, and even less on wheat and cotton yield. Some of the early studies included wheat and cotton in the analysis of increasing crop yield trends. Bean (1967) summarized an overall view of the yield trend for 18 different types of field crop including corn, soybeans, wheat, and cotton. Considering corn as the dominant field crop, Bean looked at the effect of weather and technology on Iowa corn yields and U.S. corn yields as a whole. A simplified graphical model was used to observe technological effects by holding the weather variables constant. This was done by choosing the years when weather was most favorable and least favorable based on the yield records. Bean concluded that the analysis for corn could be applied to the other crops as well as other states.

Studies by Garcia et al. (1987) and Menz and Pardey (1983) found that the increasing trend in corn yields in the U.S. was primarily due to the adoption of new agricultural technologies. Schroder et al. (1984) noted that any issues related to the

contribution of specific technologies to changes in agriculture production assume something about the underlying production function.

Swanson and Nyankori (1979) assessed the impact of weather and technology on yield growth of corn and soybeans on the Allerton Trust Farm in Piatt County, Illinois for 1950-1976 by comparing yield trends not adjusted for weather with yield trend adjusted for weather. They used monthly temperature and precipitation data for June, July, and August. Their analysis showed that yield increases follow a linear time trend, which serves as a proxy for technology; they found that using various non-linear formulations did not significantly improve the model.

Huff and Neill (1982) expressed yield as a function of time and weather variables in their study of corn yield for regions of the Midwest for 1931-1975. They concluded that July and August temperature and July precipitation are the most important explanatory variables. This corresponds to the relatively short reproductive stage (grain formation period), a two to three week period in July in the Midwest, and the historical fact that favorable August weather can enhance yield. They found the quadratic trend (including both linear and quadratic time terms) as statistically adequate to represent technological improvements.

Thompson (1969, 1970, 1985, and 1988) examined the relationship between technology, monthly weather variables, and crop yields. His findings suggest that high precipitation during July was favorable for corn and soybean yields. Linear and quadratic time trend proxies were used to represent technological change for 1960 onwards, and the result suggested that technology was not solely responsible for the increased crop yields.

Garcia et al. (1987) examined the relationship between yield level and yield stability, advances in technology, and weather conditions for corn. They divided the yield data for 1931-1982 into two different sets based on the history of technological advances. Using a linear time trend as a proxy for technological advances, they found that yield behavior adjusted for weather resulted in nearly identical yield variances for two different periods (1931-1960 and 1961-1982), which suggests that technology is not the only determining factor responsible for yield behavior.

Kaufmann and Snell (1997) estimated a hybrid model accounting for both climate and social determinants of corn yield using data from counties in the eight largest corn producing states for 1969-1987. County level data captured the significant variations in temperature and rainfall occurring within the states. They used a time trend to represent the effect of technological advances and hybrids that could not be clearly measured in other way.

Hu and Buyanovsky (2003) investigated the climate effects on corn yield data from Sanborn Field in Columbia, Missouri for 1895-1998. The results indicated that the climate effects could be better explained by within-season variations in temperature and precipitation rather than by average growing season conditions. More recently, Schlenker and Roberts (2006) employed a reduced-form model to relate weather and corn yield using detailed daily weather records for about 800 counties in the eastern U. S. for 1950-2004. Their results indicate a significant nonlinear relationship between corn yields and temperature. Yield was found to increase with moderate temperatures but the response was not favorable after temperatures exceed 30° C.

Some of the early studies on wheat focused on showing the affect of weather on wheat yield. Zink (1940) studied the relation of weather factors to wheat yield for Levan Ridge, Utah. The study suggested that the highest correlations were with evaporation, precipitation, and the length of drought periods. The study also suggested low correlations with mean temperature, but higher correlations with minimum and maximum temperatures. Brown (1959) examined the relation of weather variables to the winter wheat yield in Box Elder County, Utah. His results showed that adequate precipitation in September, October, May, and June was important to improve winter wheat yield.

Buller (1972) studied the influence of research and policy on crop yields, mainly for wheat, corn, soybeans, and grain sorghum in Kansas for 1932-1965. He used a drought severity index and a moisture departure measure in the western region, a rainfall and precipitation measure in the central region, and an evapo-transpiration measure in the eastern region as the weather variables. Fertilizer data was estimated using agricultural census data from 1954, 1959, and 1964. He also used a time variable as a proxy to estimate effects of technology on crop yield. Overall results indicated increasing yield per acre trends for all crops studied. However, the estimated annual increase in wheat yield was less for the central region than the western and the eastern region. The results also indicated that agricultural research and production policies increased yield variability for wheat and grain sorghum in the eastern and central region, but did not lessen year to year yield variability for any crop in any region.

Manogaran (1981) developed a crop-climate-technology model to examine the effects of climate and technology on winter wheat production in eleven counties in Kansas for 1921-1977. He used pre-season (August to October) and April soil moisture deficit

measures; winter (November to February), May and June precipitation measures; and March temperatures as weather variables. The relationship between the yield and technology was assumed to be linear and a time trend was used as a proxy variable for technology. The effect of technological advances on crop yield was represented in four steps, each step with different time periods, for 1921-1945, 1946-1955, 1956-1960, and 1961-1977. The yield model was expressed as a linear function of weather variables and technology variables. The results indicated that technological advances has improved crop yields under conditions of climatic pressure as a whole, but has not been able to prevail over or reverse the impacts of adverse climate.

Reddy and Baker (1990) used the GOSSYM cotton simulation model to analyze the effects of the weather on cotton yields. The weather factors were incorporated in the simulation model for five different locations to determine if there exists a significant trend, which could capture the change in yield. The output from the simulation model indicated that weather effects on lint yield trends were neutral across the entire U.S. cotton belt.

2.5. SUMMARY

This chapter provides reviews and summaries of a wide variety of previous studies and provides the basis for the development of further research methodology. The literature review finds that little research has been conducted to support estimating the benefits of the provision of soil information. Most research is based on hypothetical assumptions or based on limited samples from experimental sites.

None of the past crop yield models identified have included soil information in their analysis. The effect of the provision of soil information can be examined by incorporating the availability of soil information in crop yield models.

CHAPTER 3

THEORETICAL BACKGROUND – VALUATION OF SOILS INFORMATION

3.1. INTRODUCTION

This section provides a brief introduction to soil survey information, the benefits derived from soil information, the theoretical framework of the value of information, and discussion of the valuation of soil survey information as a public good to provide a background and develop a theoretical foundation for the analysis that follows.

3.2. SOIL INFORMATION

The soil survey program was formally initiated in 1899 with the first report of field operations, USDA Report 64, published by the USDA Division of Soils (Smith, 1998). The soil survey program initially concentrated on the capabilities of land for agriculture production. However, many nonagricultural users also came to understand the value of soil survey information in later years. Soil surveys were improved, extended, and new classification system was developed during the 1950s (Durana and Helms, 2002). Soil surveys completed after 1950 are considered modern soil surveys.

Soil survey information is essential information needed and used by government agencies and others to make land-use decisions such as development, taxation, agricultural use, and natural resources protection (Durana and Helms, 2002). Soil surveys provide information that allows users to predict the consequences of alternative uses. Young (1973) stated that the primary purpose of soil surveys is to help make land-use decisions .

Likewise Bie et al. (1973) pointed out that producers receive optimal returns when land use and management are adapted appropriately to local soil conditions.

A partial list of the users of soil survey information includes farmers, foresters, ranchers, researchers, planning agencies, engineers, development organizations, and private investors (Figure 3.1). Farmers use soil information to manage, expand, and select appropriate farming techniques. Foresters use soil information to select sites for plantations, select tree species which vary in productivity by soil characteristics, and for other management activities. Unlike farmers and foresters, planning agencies focus on broader uses, such as agriculture to urban land and grazing land to forest land conversions. Engineers use soil information to evaluate construction sites, plan road alignments, design building foundations, and evaluate sewage disposal potential.

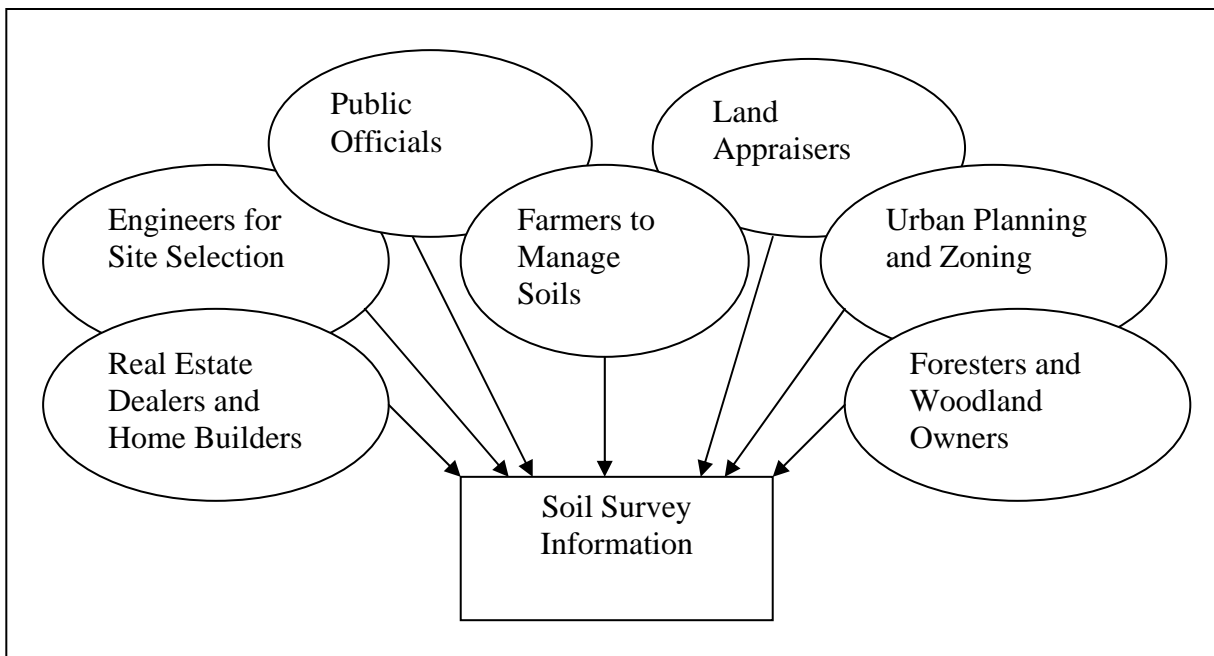


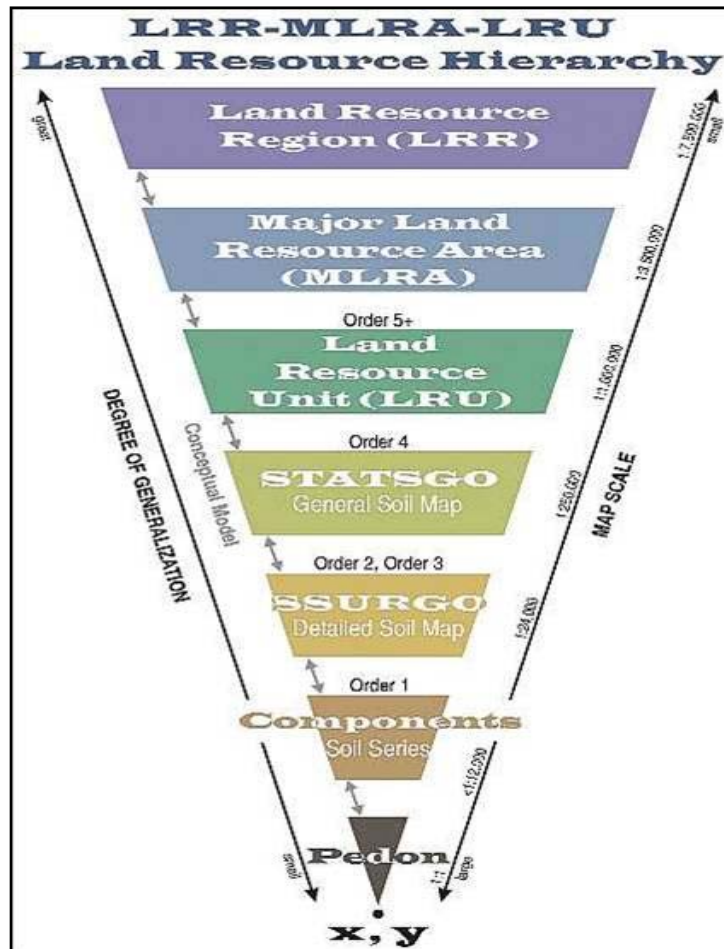
Figure 3.1: Users of Soil Survey Information

Thus the benefits derived from the NCSS program are diverse and dispersed spatially, temporarily and among user groups. Some of the benefits are immediate, some

are realized over time, and some are only realized over a long period. Aggregating economic values for a program that provides such varied and diverse benefits is complex.

Soil information has traditionally been presented as maps showing the distribution of soils in a particular area and tables that provide soil properties. Properties used in classifying soils include, but are not limited to, soil texture (grain size, color), organic matter content, moisture content, permeability, slope, elevation and, water holding capacity. Some soil information based on a one-time sample is valid for many years, e.g. elevation, landscape position, texture, and density. In contrast, regularly sampled data that reflects temporally varying information includes characteristics such as moisture content, ground water level, soil acidity, nitrogen, phosphorous, and potassium content. Soils with similar properties are grouped in mapping units.

Soil surveys are classified into five orders from the first to the fifth based on the intensity of field study, the degree of mapping detail, the phase or levels of abstraction in defining and naming map units, and different map unit designs (Soil Survey Division Staff, 1993). Figure 3.2 represents the soil geography hierarchy as a reverse pyramid proceeding from the most general at the top to the most specific at the bottom.

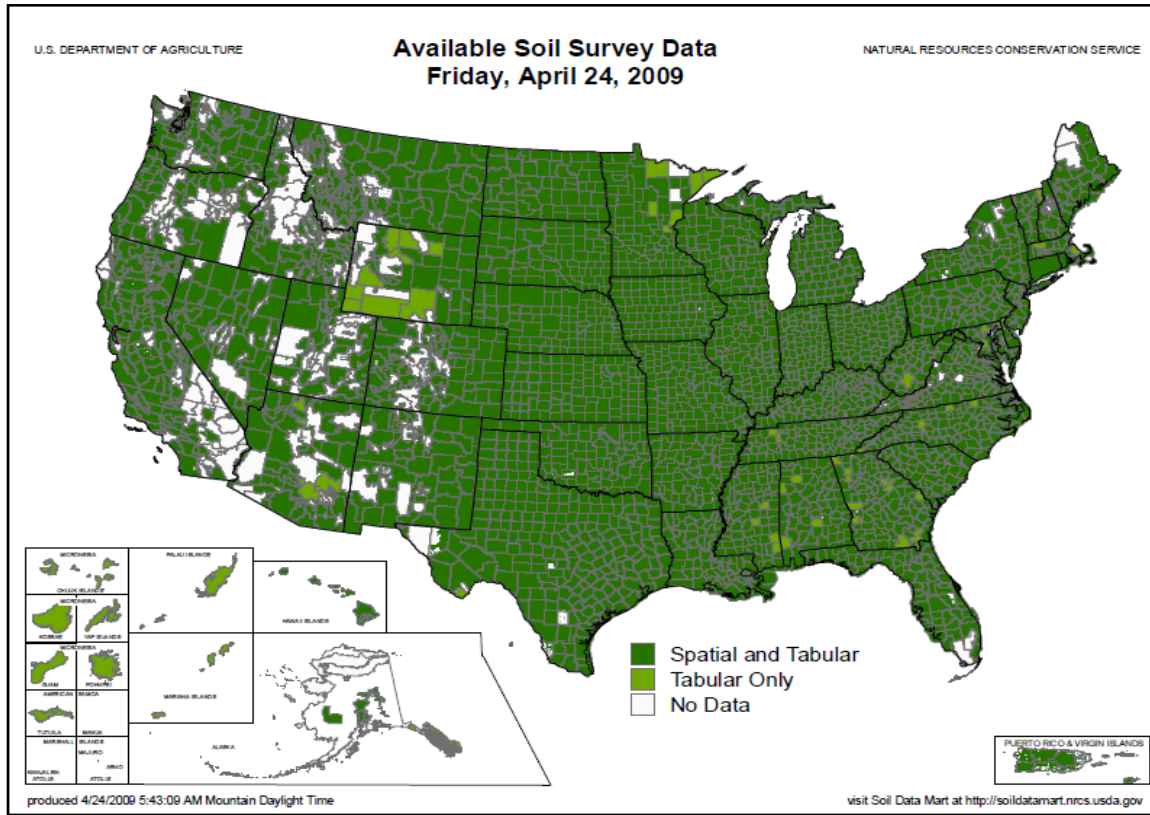


(Source: Soil Survey Division Staff, 1993)

Figure 3.2: Soil Geography Hierarchy Diagram

A first-order soil survey, the most intensive, is designed for very intensive land use planning that requires very detailed information about soils. The delineations have a minimum size of about 1 hectare (2.5 acres) or less, depending on the map scale; map scales of 1:15,000 or larger are commonly used. A second-order survey is designed for intensive land uses requiring detailed information about soil resources for predicting land suitability, use, and treatment needs. The NCSS program supports second-order surveys that are nearly complete for all private lands in the U.S. and represent cooperative efforts between state and county governments and the USDA/Natural Resource Conservation

Service (See Figure 3.3). This study evaluates the benefits of the 2nd order soil survey.



(Source: USDA-NRCS, 2009. Available online at <http://soildatamart.nrcs.usda.gov/StatusMaps/SoilDataAvailabilityMap.pdf>)

Figure 3.3: Status Map of the U.S. Second-Order Soil Survey

In summary, soil survey information developed by the NCSS program provides a detailed report on the soils for a specified particular area for use by farmers, ranchers, foresters, real estate agents, land use planners, engineers, and other organizations and individuals as well who desire information about the soil characteristics and its response (Figure 3.4).

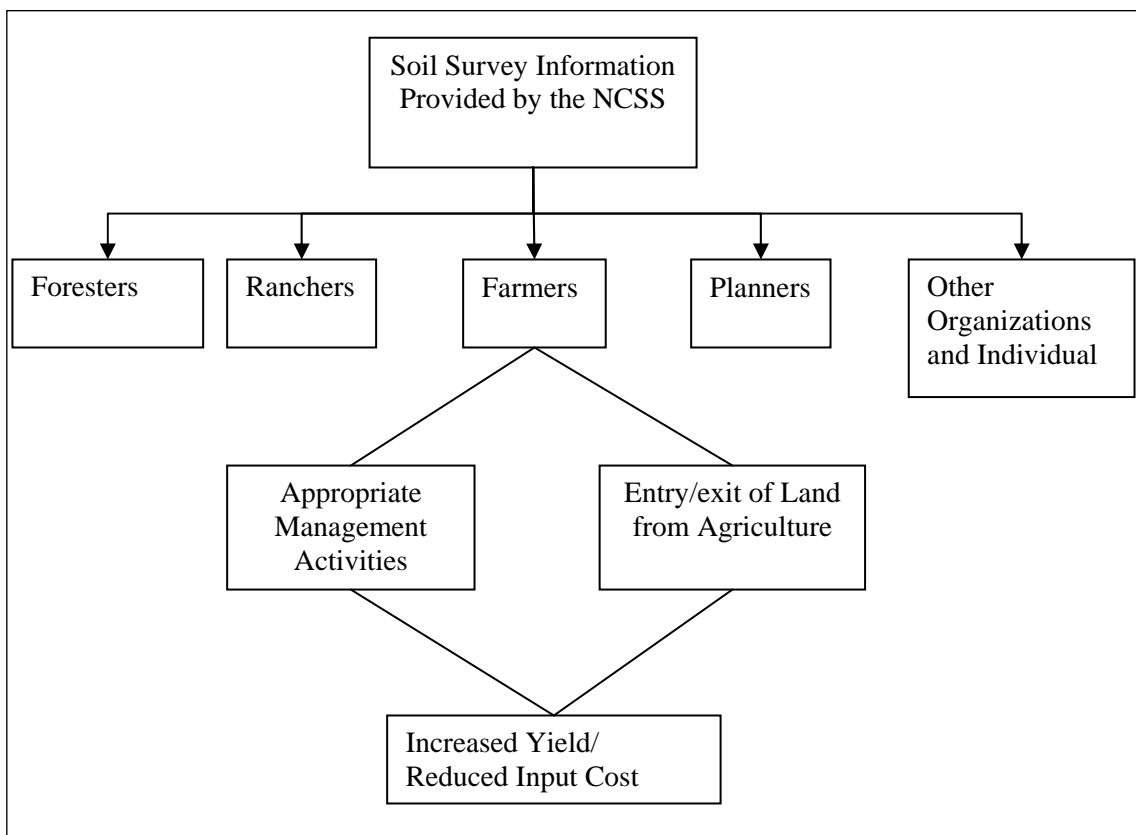


Figure 3.4: Flow Diagram Showing the Users of Soil Survey Information

3.2.1. Development of Second Order Soil Survey

Second order soil surveys are usually conducted at the county level but some may cover multi counties, only a part of a county, or parts of multiple counties. The boundaries of soil survey areas are determined at the state level with consultation from cooperating agencies of the NCSS and major users. Appropriate shape and size of soil survey areas are selected for efficient field operations and publication. Each soil survey area is named uniquely within that state, and this name is used on all records including publication reports.

The priorities for conducting soil surveys within a state are determined in consultation with cooperators and the state conservationist (Soil Survey Staff, Soil Conservation Service, 1993). Priority lists are influenced by several factors:

- requests for soil surveys from local people,
- state and local needs for information that helps in land use planning and decisions,
- state and local needs for tax evaluation,
- intense land use changes in areas facing soil problems, and
- state and local contributions in terms of both fund and manpower.

A memorandum of understanding is prepared for each soil survey area as soon as possible after the decision is made to conduct a specific soil survey (Soil Survey Staff, Soil Conservation Service, 1993). The memorandum of understanding addresses the objectives and specifications of the survey and provides a description of the area to be included. A preliminary field study is then conducted to get the project personnel familiar with the survey area. After completion of the preliminary field survey, the project leader of the soil survey along with soil scientists and other related experts confirm that the memorandum of understanding adequately explains the necessary details of the soil survey project including the purpose of the project, specifications, description of the work area, cooperating agencies, and responsibilities.

Once the memorandum of understanding is processed, field sheets are prepared from rectified photobase maps or orthophoto base maps. Each field sheet contains information on the name of the agency, the acreage of the soil survey area, the name of the soil survey area and state, map scales, name of the soil scientist(s), and the completion date.

The soil scientists request access to private lands. Once access is granted, soil scientists walk across the land observing and documenting landscape characteristics such

as slope, vegetation, aspect, parent materials, and other features affecting soil use. They dig holes to expose soil profiles and determine physical and chemical characteristics for the horizons exposed in the soil profile. Soil scientists collect data to quantify, compile to develop soil map units, and document soil performance for the soil interpretations identified in the memorandum of understanding.

The soil survey information gathered by the soil scientist is developed as a soil survey database. Soil survey systems are fundamental part of the collection, storage, manipulation, and dissemination of the soil information (Soil Survey Staff, Soil Conservation Service, 1993). The soil survey information includes description of the soils and their locations. The soil survey describes and classifies soils and contains soil interpretations appropriate for planning and discussion of the suitability, limitations, and management of the soils for specific uses.

A soil survey manuscript is prepared to facilitate the dissemination of soil information to decision makers using soil information. This manuscript is reviewed by the state soil scientist and other staff. After the soil survey manuscript is finalized, it is sent to the Government Printing Office for publication. Copies of published soil surveys are sent to depository libraries that have requested them. For others, soil information is made available upon request as a hard copy and/or electronically on the Web Soil Survey at <http://websoilsurvey.nrcs.usda.gov/app> (Soil Survey Staff, Natural Resources Conservation Service, accessed 2009).

The expected time frame to complete a soil survey project for an individual area is five years, in some cases it might take longer. After the completion of fieldwork, soil names and descriptions are correlated and approved. The approval date is called the

correlation date. Once the soil survey is approved, the soil survey information compiled from the survey is sent for publication. Due to delays in publication from technical, budget, or other issues, the gap between the correlation and publication date can range from less than a year to several years. For example, major fieldwork for Brown County, Illinois was completed in 1982. Soil names and descriptions were approved in 1983 and the report was published in 1988 (USDA-Soil Conservation Service, 1988). Similarly, major fieldwork for Ozark County, Missouri was completed in 1999, soil names and description were approved in 2000, and the report was published in 2000 (USDA-NRCS, 2000).

However, soil survey information is commonly provided before the publication of the soil survey report. Soil information is often made available by the field crew during the time of fieldwork and preliminary data is commonly available after the correlation date. Those with a knowledge of soil science can use soils information to predict the response of specific soils to various uses and management activities.

Maintaining soil survey information is an ongoing activity. The purpose of soil survey maintenance is to provide current, accurate soil information to users and often to add additional information from advances in the underlying soil science discipline. If the published soil survey is outdated, inadequate, and deficient and appropriate resources are available, a memorandum of understanding is prepared for the maintenance and corrective measures are taken. Soil surveys are updated thorough continuing data collection, regular reviews, assessment, and additions to existing soil survey information. Thus many counties have more than one soil survey with the later soil surveys providing additional soil information for a boarder scope of uses.

3.3. VALUE OF INFORMATION

According to information theory, information is defined as the reduction in uncertainty. McGee and Prusak (1993) defined information as data, both factual and numerical, that is organized and imbued with meaning as a result of gathering, analyzing, or summarizing the data in a meaningful way. Data are considered as outcomes of query processes involving sampling or from experiments. For example, soil data result from field samples and measurement augmented with laboratory analysis. The supply of information depends on the production cost and challenges of collection. Demand for information and its value depends on its role for improving economic decisions.

Economic principles consider information valuable if it leads to preferred decisions (Preckel, Loehman, and Kaylean, 1987). Information is then valued by the difference between outcomes obtained with the information and without the information. The value of information is an outcome of choice in uncertain conditions (McCall, 1982). It is the difference between the project value with the information and the project value without the information, minus the cost of acquiring the information. The value of information is determined by its importance to the decision makers or to the outcome of the decision. Decision makers may be willing to pay for information depending on the degree of uncertainty and what is at stake (Macauley, 2005).

Information has value when the alternative outcomes can be different; otherwise information has no role in adding value. In other words, there must be uncertainty, and if there is uncertainty, there must be choices. If there are no choices, there are no decisions to be made, and information has no value. Thus information is considered valuable if it leads to a preferred decision. In the decision theory literature, the value of information is

defined as the difference in expected utility (in terms of the appropriate probability distributions) between the decisions made with more information and less information (Preckel, Loehman, and Kaylen, 1987).

More information helps individuals make better decisions. Information helps decision makers in a variety of ways that lead to better decisions such as identifying the problem, developing and evaluating alternatives, and selecting and implementing the best alternatives. Better decisions increase expected utility. Individuals are expected to be willing to pay for additional and improved information if the cost of the information is lower than the expected value of their gains.

Macauley (2005) specifies that the value of information depends on the following factors:

1. Degree of uncertainty of the decision maker: How much will information help in making the decision? If there are few actions available, information can have low value.
2. What is at stake (value of the alternative outcomes of the decision): Value of the outcome is the total value of resources or activities as an outcome of the decision. Willingness to pay for information is a derived demand. How much could the final value of the outcome be affected?
3. Cost of information used to make the decision.
4. Price of substitutes for the information: Are there any alternatives? If so, what is the cost for the substitute?

The larger the degree of uncertainty and the value of output, the larger the value of information; the larger the cost of information and the lower the price of substitutes, the smaller the value of information.

3.3.1. Theoretical Framework:

Theoretical aspects of the value of information have been discussed by Lawrence (1999), Hilton (1981), and Radner and Stiglitz (1984). The theoretical background discussed in this section is adopted from Lawrence (1999). According to Lawrence (1999), the value of information is the difference between the expected payoff incorporating information and the expected payoff without incorporating information expressed as

$$(3.1) \quad V(x, y) = \pi(x, a_y) - \pi(x, a_0)$$

where, x is the realized state,

y is the information,

a_y is the optimal action to the information y under present knowledge $p(x|y)$,

a_0 is the action without information y , and

$\pi(x, a)$ is the payoff.

The value of information can be positive, negative, or zero. If incorporation of information in the decision improves the outcomes, it has a positive value. If the incorporation of information in the decision reduces the net outcomes than it would have been achieved without the information, then it has negative value¹. If the incorporation of information in decision has no effect in outcomes, then the information has zero value. To

¹ This can occur in only two situations, the information is either wrong or overwhelming. Too much information inhibits decision maker's choice.

access the net impact of information (NI) when outcomes are certain, the cost of information should also be incorporated which is expressed as

$$(3.2) \quad NI = V(x, y) - c = \pi(x, a_y) - \pi(x, a_0) - c$$

where c is cost of accessing information.

However, it is decision maker's choice whether or not to incorporate information into decision. It depends on the cost of the information and the potential for the information to improve the outcome. If the cost of information outweighs the expected benefits from using information, a rational decision maker will not use the information.

3.3.1.1. Expected value of Information

Equations (3.1) and (3.2) can be used to estimate the ex-post value of information because outcomes are known. However, there are analytical difficulties to estimate ex-ante value of information due to uncertainty. In the case of uncertainty, decision makers or the system designers need a criterion to evaluate and compare the possible alternatives. The criterion of maximizing expected value is the most commonly used basis to evaluate and compare alternative options.

Using this expected value criterion, the optimal choice is identified by computing the expectation of the payoff function $\pi(x, a)$ for every action $a, a \in \mathcal{A}$, and choosing the action a_0 that maximizes the payoff,

$$(3.3) \quad \max_a \int_X \pi(x, a)p(x)dx = \max_a E_x \pi(x, a) = E_x \pi(x, a_0)$$

where $p(x)$ = initial knowledge. Similarly, if the decision maker incorporates the information in her system $p(x|y)$, the optimal action after incorporating the information y is given by:

$$(3.4) \quad \max_a \int_X \pi(x, a) p(x|y) dx = \max_a E_{x|y} \pi(x, a) = E_{x|y} \pi(x, a_y)$$

Where $E_{x|y}$ is the expectation with respect to $p(x|y)$. The action a_y is the conditional decision rule that informs the decision maker what to do, conditional upon the addition of new information y into the new state of knowledge expressed by $p(x|y)$. $E_{x|y} \pi(x, a_y)$ gives the decision maker's expected payoff. So now, the value of information is expressed by the difference between the expected values of payoff or outcome.

$$(3.5) \quad V(I) = E_{x|y} \pi(x, a_y) - E_x \pi(x, a_0)$$

3.3.1.2. The Utility Function

The terminal level of the decision maker's wealth is sometimes given more importance than just the payoff, $\pi(x, a)$, in decision problems with quantifiable outcome in monetary units. In this case the outcome is change in decision maker's total wealth, including fixed and known initial wealth w to the terminal wealth W . Suppose W represents the set of decision maker's potential terminal wealth, $W \in W$. Now the outcome function ω is expressed as function of w, x , and a .

$$(3.6) \quad W = \omega(w, x, a)$$

Suppose the payoff from the decision problem is additively separable from the initial wealth, the terminal wealth outcome can be expressed by

$$(3.7) \quad W = w + \pi(x, a).$$

Utilities are used when the decision criteria is based on more than the expected payoff, $\pi(x, a)$ or the terminal monetary wealth. Utility is measured as the total value of a particular outcome. The decision maker's utility function, defined on W , can be expressed as

$$(3.8) \quad U(W) = U(\omega(w, x, a)).$$

Utility function $U(W)$ is assumed strictly increasing and continuous in W . The optimal decision can be determined using the expected utility approach. The expected utility of a decision D , conditional on initial knowledge I_i and an action $a \in \mathcal{A}$, is represented as

$$(3.9) \quad U(D | I_i, a) = E_x U(\omega(w, x, a)).$$

The optimal prior decision that is without information y , is the choice of action a_0 that maximizes the utility of decision maker and can be represented by

$$(3.10) \quad \begin{aligned} U(D^* | I_i) &= \max_a U(D | I_i, a) \\ &= \max_a E_x U(\omega(w, x, a)) \\ &= E_x U(\omega(w, x, a_0^*)) \end{aligned}$$

where, $U(D^* | I_i)$ is the value of prior decision. Similarly, optimal decision with information y can be derived by

$$(3.11) \quad \begin{aligned} U(D^* | I, y) &= \max_a E_{x|y} U(\omega(w, x, a)) \\ &= E_{x|y} U(\omega(w, x, a_y^*)). \end{aligned}$$

The expression $U(D^* | I, y)$ represents the decision maker's before-cost utility prior to realization of the state but after the incorporation of information.

Lawrence (1999) employed a conditional decision rule, and combined with the above equations showed that for an expected-utility-maximizing decision maker, the value of informed decision is at least not less than the value of prior decision.

$$(3.12) \quad \begin{aligned} U(D^* | I) &= E_y \max_a E_{x|y} U(\omega(w, x, a)) \\ &\geq E_x U(\omega(w, x, a_0^u)) = U(D^* | I_i). \end{aligned}$$

Left hand side of the equation (3.12) represents maximum conditional expected utilities given all possible sets of information (represented by E_y) and the right hand side of the equation represents maximum utility without information. Equation (3.12) shows that expected-utility-maximizing decision maker will be as well off by incorporating cost-free information compared to prior decision without that information. Thus it can be concluded that if farmers act as expected-utility-decision makers, they would not be made worse off by using cost-free soil information from the NCSS.

3.3.1.3. Stages of Valuing Information

According to Lawrence, the value of information could be measured at any of four stages:

1. Prior or Ex-ante
2. Ex-post
3. Conditional
4. Pre-posterior

Prior or ex-ante value of information is the expected value before taking into consideration of incorporating information. Ex-post value of information is value posterior to both application of information and realization of the state. Conditional value of information is the posterior value to application of information but before realization. The pre-posterior value is the value before processing information.

3.4. VALUE OF SOIL INFORMATION

The information provided by the NCSS has played a significant and important role in diverse fields. There are considerable challenges to estimate fully the aggregate benefits derived, but such estimates are needed to conduct an accurate cost-benefit analysis.

Temporally, benefits provided by soil surveys can be broadly categorized into the following three groups:

1. Historical benefits
2. Current benefits
3. Future benefits

Historical benefits are the benefits achieved in the past period from the use of NCSS developed soil survey information, current benefits are the benefits realized in the current period from the use of NCSS developed soil survey information, and future benefits are benefits expected in future years from the availability of NCSS soil survey products as well as continuing activities. Historical and current benefits are derived from past investments. Future benefits can be further divided in two categories: 1) benefits to be derived from past investments in the NCSS program, and 2) benefits that will be derived from additional (current and future) investments in the NCSS soil survey program.

Estimates of past benefits provide a measure of the returns to past investments. Current benefits give a measure of the ongoing returns to past investments. All types of benefit estimates depend on time, duration, uses, and the user groups considered. Some of the benefits realized in the past and continuing in the current period can be estimated through indirect methods using currently available data. Partial future benefits can be estimated by extrapolating from such analyses.

3.4.1. Benefits of Soil Survey Information in Agriculture Production

The primary goal of the soil survey program is to assist society and individuals to understand the suitability and limitations of the soil resources for intended uses (Ditzler, Engel, and Ahrens, 2003). Soil information has been used for centuries to guide farmers to manage and better understand crop growth (Samuelson et al., 2002). Soil maps and the attributes of the various soil series derived from the soil survey provide information to farmers for site selection, land use, and management activities. Farmers use soil information to determine the capability of soils to sustain certain kinds of crops, the relative productivity of farm fields, and the best production practices for a given situation. Soil information can thus affect agricultural production related decisions on both the intensive and extensive margins. An increase in aggregate supply can result from adjustments on both the intensive and extensive margins. Thus provision of soil information can be expected to change the supply curve for a particular crop.

The intensive margin refers to the degree or intensity of how a resource is utilized and managed. On the intensive margin, soil information affects crop and rotation choice as well as fertility, tillage, and other management activities. Introduction of soil information helps farmers to better understand and manage their land which can increase yield and/or reduce costs. It could be done by changing the level of inputs, such as need for irrigation depending on soil capability, applying fertilizers depending on the soil quality, changing management activities, and changing cropping patterns and rotations.

On the extensive margin, soil information affects land purchase decisions and stimulates movement of marginal lands in and out of production, which also affects aggregate supply curves. Even though the amount of U.S. land used for crops has remained relatively constant for the last century, a large amount of land enters or exits out

of cultivation each year. Total U.S. cropland used for crops was 330 million acres in 1910, 377 million acres in 1950, and back to 330 million acres in 2006 (USDA-ERS website, 2008). During the 1982-1997 period, 60 million acres of cropland shifted to less intensive land-use such as CRP grazing, forestry, and other rural uses. Likewise, 26 million acres of less intensive land-use was changed to crop land in the same time period, and 12 million acres moved from uncultivated cropland to cultivated cropland (Lubowski et al., 2006). In the 1979-1981 period, 2.2 million acres of land was converted to cropland in the Corn Belt region (Heimlich, 1986). Soil quality along with other variables such as scale of production, government policies, and other factors affecting the relative profitability of growing crops plays a major role in determining the maintenance of cultivated cropland at the margin (Lubowski et al., 2006). Information on aggregate soil quality is primarily obtained from soil survey reports developed by the NCSS program. The NCSS program thus plays a significant role in the movement of marginal lands in and out of crop production.

3.4.1.1. Hypothetical Case

Soil surveys classify land as agricultural or non-agricultural. For agricultural land, soil survey reports provide information on the suitability of the land resource for specific crops, thus provision of soil information helps to increase the utility of farmers. The value of soil information results from the farmer's increase in utility from using soil information to improve decisions. Take a simple example: let us assume a farmer has two types of land based on soil properties, land 'A' suitable for crop a, and land 'B' suitable for crop 'b' given current production practices. However, land 'B' could be suitable for crop 'a' if management practices are changed. The farmer derives utility of 1 if he plants the right

crop or appropriate management activities based on the soil, otherwise less than 1. Soil survey information helps farmer to distinguish between the responses of his land to different crops under alternative practices (Figure 3.5). Thus the provision of soil information in combination with other factors could help the farmer to increase utility.

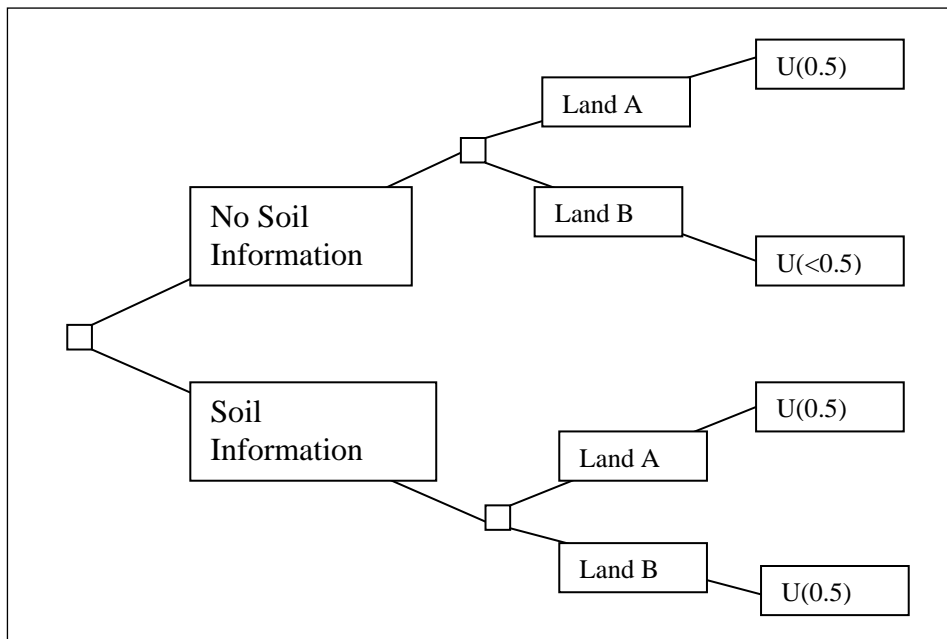
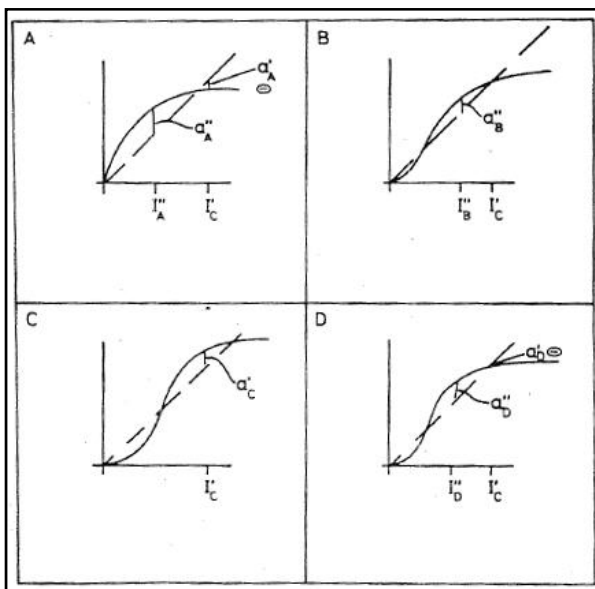


Figure 3.5: Simplified Decision Tree for Farmer with Land ‘A’ and Land ‘B’ (Assuming Soil Survey is Perfect Information)

Beckett (1981) discussed a similar hypothetical case where the farmer wanted to optimize his profit by using all of his land for the same crop. The farm land was composed four different types of soil, each with a different capacity to produce that that specific crop. Without the soil survey information, the farmer had assumed his land was of a uniform soil type with the same response to management and inputs for that crop throughout the farm. He applied inputs on all of his lands that were appropriate for only one of the four soil types. For example, the profit would not be optimal if the same amount of fertilizer was applied to different types of land. A soil map could provide the

information necessary to predict the responses for the specific crop and could have led to different levels of inputs depending on the soil response. A simple soil survey map could have helped the farmer to increase his payoff and/or reduce the cost of fertilizer. Figure 3.5 shows a farm with four soils A, B, C, and D of equal area. Each soil had significantly different response curves for a particular crop. The optimum input for soils are represented by I' and I'' respectively, for the first case (without soil information) and the second case (with soil information). In the first case, farmer applied I'_c to all of his land assuming that all of land lies in soil C. In the second case, after realizing the fact that the lands were in different type of soils, he applied inputs ($I''_A, I''_B, I''_C,$ and I''_D) according the soil type that would optimize the output. Optimal profits are represented by $a''_A, a''_B, a''_C,$ and a''_D in Figure 3.6.



(Source: Becket, 1981)

Figure 3.6: A Farm Lying on Four Soil Types (A-D), with Different Response Curves for a Particular Crop

3.4.2. Economic Analysis of Availability of Soil Information

Policies having nonprice effects on the producer must sometimes be evaluated (Just, Hueth, and Schmitz, 2004). The government has made investments in collecting and providing soil information for the public good. This information, provided to users free of charge, substantially affects aggregate productivity (Just, Hueth, and Schmitz, 2004). To account for nonprice impacts on producers, the interpretation of fixed factors of production can be expanded. Since such factors do not exist in markets, demand for such factors is not directly observable.

Benefits derived from soil survey information can be demonstrated using a standard supply and demand framework and economic welfare methods. Supply of any good depends on price and production cost as well as other factors. The introduction of soil information may change the supply curve for a particular crop. An increase in supply can result from adjustments on both the intensive and extensive margins. Introduction of soil information helps farmers to better understand and manage their land which can increase yield and/or decrease costs. Soil map provides information to the farmer that is required for different management to optimize returns (Bie et. al, 1973). Thus soil information helps to improve farm efficiency. Soil information also affects land purchase decisions and stimulates movement of marginal lands in and out of production, which also affect crop supply curves.

The benefits derived from the information provided by soil surveys and the cost to produce the soil surveys can be computed using generally accepted welfare economics methods. Welfare economics is based on the idea that a change in an individual's economic well-being can be measured in terms of the individual's willingness to pay to obtain the change (in case of a good) or willingness to pay to avoid (in case of a bad). All

individuals in society are categorized as producers, consumers, or both in order to analyze changes in social welfare in market terms. Consumers' welfare is measured by consumer surplus (as a first approximation) while producers' welfare is measured by producer surplus. In Figure 3.6, suppose D is the demand curve and S_0 the initial supply curve for a crop (e.g. corn). The area below the demand curve and above the initial price, P_0 , bounded by the initial supply curve, S_0 , represents consumer surplus (area ABP_0). Producer surplus is the area above the supply curve and below the price line, P_0 , (area P_0BC).

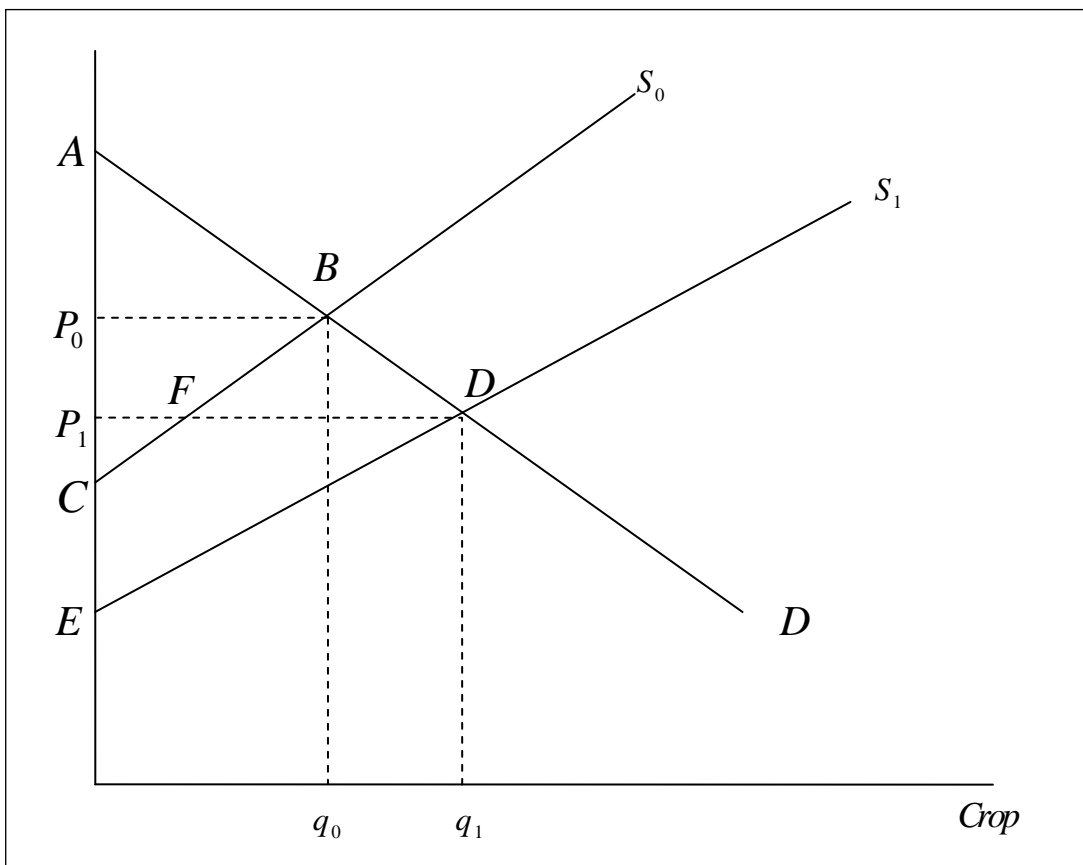


Figure 3.7: Welfare Analysis in Market Equilibrium Framework

Assuming the change in supply of the crop due to provision of soil information reflects true social value, the welfare effects are represented by Figure 3.7. The initial equilibrium, i.e., before the availability of soil information, is represented by point B , the point that generates maximum social welfare, i.e., the sum of producer and consumer surplus when farmers do not have soil information. The use of soil survey information helps farmers better manage their land and affects decisions regarding the inclusion of land in crop production. Thus introduction of soil survey information increases yield and reduces the marginal and average costs of production, shifting the supply curve outwards from S_0 to S_1 ; this results in a new equilibrium at point D . This results in an overall gain in social welfare equal to the area $BDEC$. Because of the reduction in price due to higher output, consumers unambiguously gain an amount equal to area P_0BDP_1 . Producers gain an area of P_1DE less P_0BFP . The producers' net gain from the introduction of soil survey information is ambiguous, depending on the relative elasticity of supply and demand. If the demand is elastic, producers are likely to gain. However, if the demand is more inelastic, producers are likely to lose.

3.4.3. Soil Information as a Public Good

According to economic theory, a public good is a good that, once produced, can be consumed by an additional consumer at no additional cost. Goods and services that are both nonexcludable and nonrival are public goods. Nonexcludable means that no one can be excluded when the good is provided. Nonrival means that one person's consumption does not reduce the ability of other to consume that good. Public goods are jointly provided, and the benefits accrue collectively to society. These goods are not divisible into units that are appropriated to individuals. Information goods are reasonably nonrival,

because of nominal cost to reproduce it. However, nonexcludability depends on the cost of exclusion and sometimes on the legal regime. The public good characteristics of information related to ownership and difficulty of exclusion has led to discussion among economists (Braunstien Y, 1981; Chavas and Pope, 1984).

Soil survey information in the U.S. is considered public property (Soil Survey Division Staff, 1993). Soil information, as a public good, is provided by the NCSS program through soil survey publications and web services. Thus the second order soil information provided by the NCSS program in the U.S. has the basic characteristics of a public good.²

If the individual landowner want the soil information for their land, it would be inefficient for them to conduct a soil survey to obtain the required soil information for their land only. Since such information would be privately held, it would not be available for comparison or available generally to those looking to acquire land. However, once a soil survey has been completed and the information made available to the public by the NCSS program, it can be used for any purpose and by any potential user. The soil survey program produces and maintains quality soil survey information efficiently by planning, directing, guiding, and maintaining the NCSS program at all administrative levels (Soil Survey Staff, Soil Conservation Service, 1993). For example, the size and shape of survey areas are chosen for efficiency for both field operations and publications. The minimum size of a second order soil survey area for efficient publication is about 200,000 acres while the maximum size is about 1,000,000 acres (Soil Survey Staff, Soil Conservation

² However, in some cases detailed or more accurate soil surveys are conducted by individuals to produce soil information to meet their needs. In this case, soil information is a private good produced by individuals or the private parties. For example, first order soil survey for precision farming could be considered as private good if it is produced by the farmers themselves. Likewise some of the timber companies and construction companies produce their own detailed soil survey information.

Service, 1993). Two or more small counties are often combined in a single survey area, and large counties may be subdivided into more than one survey area to obtain a size and shape that is efficient for field operations and publications.

Based on the discussion above, soil information provided by the NCSS has the characteristics of a public good. Because of the public good nature of soil survey information, the economic value of the soil survey is not directly observable in market transactions. It is thus difficult to estimate the economic value from additional investments. In these cases, non-market valuation techniques could be applied to estimate the economic value of soil information that society receive from uses of soil survey information.

There is no doubt that soil information has value and plays a significant role in decision making. However, since information is not usually traded in markets, quantifying the value of information is difficult and complex because it involves the decisions the individual would have made without information and the consequences of those decisions. Because of its subjective nature, it is difficult to quantify and value information directly. Since it is costly to produce soil information and inefficient for the individual user to produce soil information, it is important to estimate the value of such information to society. The approach suggested here for benefit estimation demonstrates an innovative approach for valuing information and provides a measure of benefits that can be used to conduct an aggregate benefit-cost analysis.

3.4.4. Nonmarket Valuation Approaches to Valuing Soil Survey Information

Because of the public good nature of soil survey information, the economic value of the soil survey is not directly observed in market transactions. In these cases, non-

market valuation techniques could be applied to estimate the economic value of soil information that society receive from uses of soil survey information. The value of public goods can be measured as willingness to pay (WTP) and willingness to accept (WTA) using direct value elicitation methods. WTP is the maximum amount of money an individual is willing to pay for the improvement (additional investment) and WTA is the minimum amount of money the individual would require to forgo the improvement (Freeman, 2003). In the case of additional investments in the soil survey, WTP is a compensating variation measure of welfare change, whereas WTA is an equivalent variation.

Two approaches could be employed to estimate the benefits provided by soil information:

- Direct methods
- Indirect methods

Direct methods are survey based approaches to valuation usually based on individual responses. Such methods attempt to determine the value for a public good by directly asking individuals. The contingent valuation approach is a commonly used direct method based on the decision maker's responses to hypothetical questions. Properly constructed, such surveys provide the information needed to conduct traditional demand analysis. It is one of the oldest methods to elicit consumers WTP for nonmarket goods (Young, 2005). Mitchell and Carson (1989) argue that the contingent valuation method is the most promising approach for determining WTP for many public goods, if the method is applied carefully.

Indirect methods involve observing real world behavior in response to a public good and then applying economic models and statistical analysis to extract and identify the value of the public good. Indirect methods rely on statistical procedures within an accepted economic framework to capture the impacts on decisions and related outcomes. The analysis of econometrically estimated production and demand functions provides an example of the use of indirect methods.

The production approach begins by trying to measure the contribution of the public good to output derived from its use through standard aggregate production relationships that depend on a vector of standard factors of production in addition to the soil information. For instance, the impact of soil survey information on aggregate corn production using a panel data approach of average county corn yield over time for several hundred counties can capture the impacts of temporally distributed access to soil survey information in aggregate production functions.

The general form of the production function expresses output as a function of a vector of factors that contribute to output in addition to the soil survey information:

$$(3.13) \quad Y = f(S, X)$$

where Y represents average crop yield, S represents soil information and X represents the vector of factors that determine crop yield trend such as technology, hybrid, weather, fertilizer, and pesticides. The effect of the introduction of soil survey information or a change in soil survey information for a county can be estimated by measuring the impact on crop yield correlated with the provision of information that is not explained by the usual inputs. The final form can be manipulated to isolate the impacts of soil information.

This study relies on indirect methods to analyze econometrically estimated production relationships to measure the value of soil information. The primary development is through a case study of corn production in the Corn Belt and extensions of the study to other major crops. The dissemination of soil survey information over the past 60 years in conjunction with the data available on corn production provide the results of a natural experiment that can be analyzed to evaluate the impacts of the NCSS on corn production.

3.5. NATURAL EXPERIMENT

A natural experiment is a naturally occurring event which facilitates the ability of a researcher to answer a specific question. Researchers are able to use natural experiments when the data from controlled experiments are difficult or impossible to obtain but some set of events have led to outcomes that can be interpreted and analyzed as if the variation in outcomes were attributable to an experimental design. Meyer (1995) stated that natural experiments can be influenced by government randomization, policy changes, or other events that provide the opportunity for a researcher to acquire exogenous variation in the main explanatory variables. Government policies often result in a set of outcomes and an environment that can be interpreted as a natural experiment. This can happen when the government policy allow changes in some states and not in some others. For example, the county level soil survey is conducted in each county in different period of time, thus providing soil information for counties at various times, some earlier and some later. This aspect of the NCSS program provides a series of outcomes that can be interpreted as a natural experiment in this study to assess the impact of the provision of soil information on crop yields. This cross-county difference in availability of soil information and the

timing of impacts provides the structure that this study uses to analyze the value of soil information in agriculture production.

However, there are limitations to the use of natural experiments. The researcher has no control over how the explanatory variables have been influenced by other non-treatment factors (Leblanc, 2004). Because of this, the data derived from natural experiments are difficult to analyze to understand the observed differences between the treatment groups caused by natural treatments.

Leblanc pointed out that in spite of the drawbacks inherent in natural experiments; there are a number of justifications for their application. Some issues cannot be easily studied in controlled experiments and the responses of experimental subjects are more realistic. Conducting an experiment for valuing soil information for a whole state or at a national level over time would be impossible. The cost could rival that of the NCSS program itself.

CHAPTER 4

EMPIRICAL MODELS AND DESCRIPTION OF DATA

4.1. INTRODUCTION

The spatial and temporal dissemination of the information provided by the NCSS program over the past 60 years in conjunction with data available on crop production provide the outcomes of a natural experiment that can be analyzed to evaluate the impacts of the NCSS program on crop production. The primary development is through a case study of corn production in the Corn Belt and is extended to selected other major crops. The innate spatial variability in crop production is captured by a county level productivity index and the temporal trend is captured in a time trend that captures a variety of technology enhancements. The primary stochastic processes that drive spatial and temporal variability are captured by county level weather measures for each year.

The crop model is based on the knowledge that primary production in agriculture is dependent on climate, soil, and the level of technology in a society. The yield of an agricultural crop is governed by the nature of the soil, weather, and management practices (Simonson, 1955). The model is based on estimating yield trends as a function of spatially and temporally varying weather data, own price, spatially variable soil productivity, time trends that reflect technical and management change, and the timing of the introduction of soils information by county as soil surveys were completed.

4.2. STATISTICAL MODEL OF CROP PRODUCTION

No past crop model has been identified that includes the variable for provision of soil information as an explanatory variable. A model integrating soil information with other variables such as technology, weather, own price, and a productivity index can be employed to estimate the contribution of soil information to aggregate crop yield. The models developed for crop yield are based on specifications provided by several previous studies (Kaufmann and Snell, 1997; Garcia et al., 1987; Schroder et al., 1984; and Thompson, 1969 1970 1986 1988). The general form of this model is expressed as:

$$(4.1) \quad \text{Yield} = f(\text{soil survey information, technology, climate, price, soil productivity})$$

The spatially wide spread and temporally diverse nature of the provision of soil survey information supports the contention that the provision of the soil survey information was not systematically correlated with other variables such as technology, fertilizer use, and the introduction of hybrids. To the extent that these assumptions hold true, the methods utilized in this research provide an unbiased estimate of the impact of soil information on crop production.

Figure 4.1 and 4.2 show the random pattern of soil survey and Figure 4.3 show the county map pattern of average corn yield for 2007. Provision of soil survey information is not highly correlated with the county size, crop acreage or corn yield (Table 4.1 and Figure 4.4 - 4.6). Within the model used in this research, spatial variability is captured by the county level productivity index and the various temporal trends are captured in a time technology trend. The primary spatial and temporal variability is captured by the county level weather measures for each year.

Table 4.1: Correlation Matrix - Provision of Soil Survey Information, Crop Acreage, and Crop Yield

	Provision of soil survey information	Average yield	Crop Acreage
Provision of soil survey information	1	-0.2371	-0.03571
Average yield	-0.2371	1	0.2642
Crop Acreage	-0.03571	0.2642	1

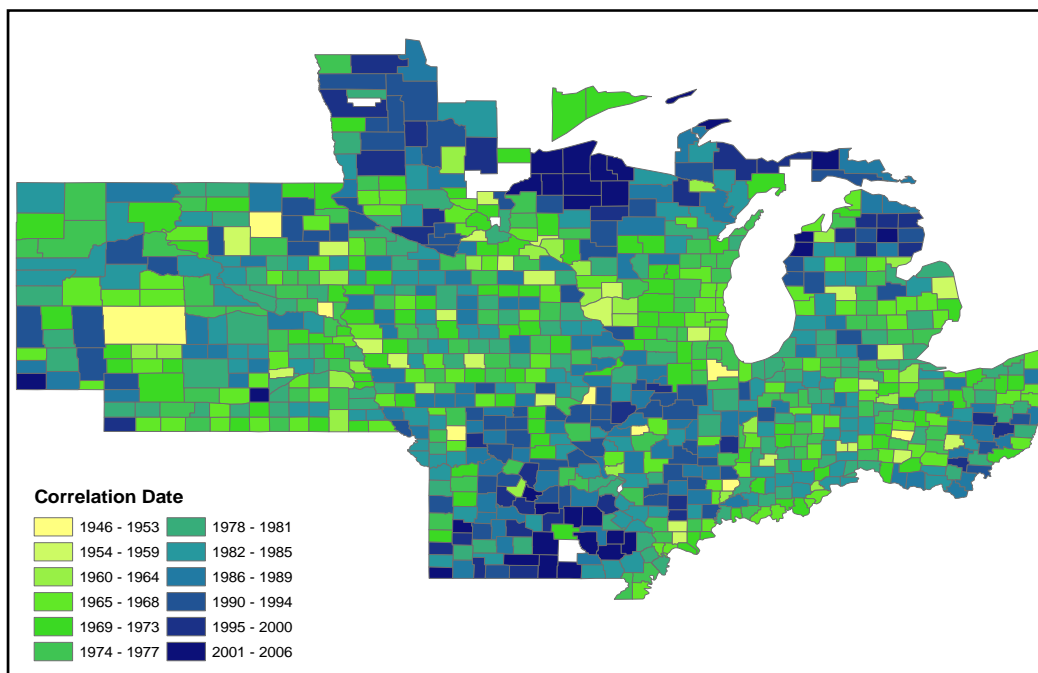


Figure 4.1: Correlation Date for Major Corn Producing Counties in the U.S.

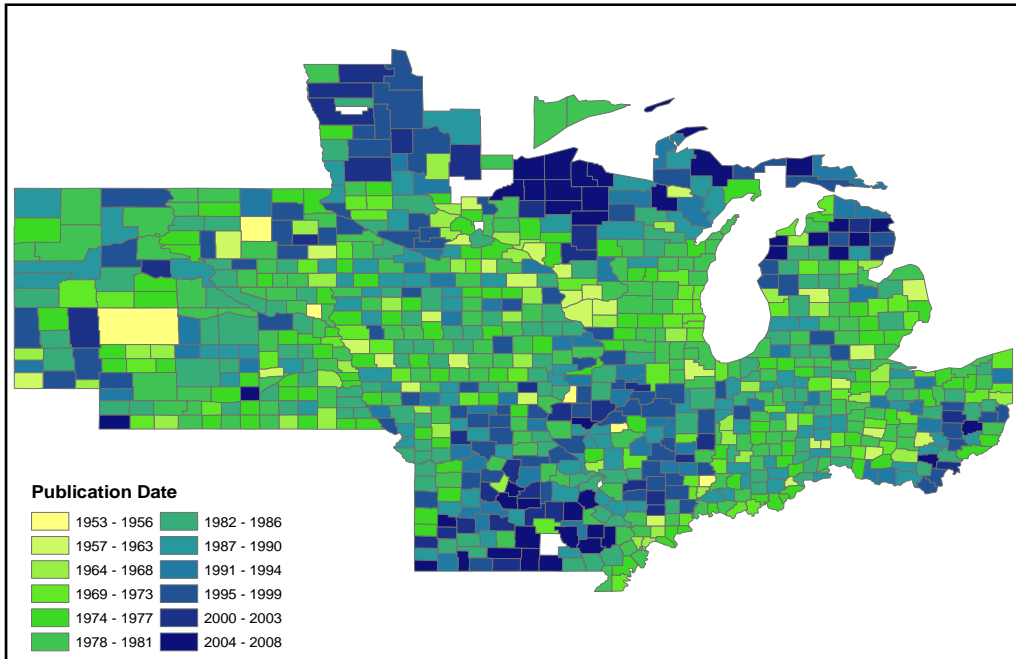


Figure 4.2: Publication Date for Major Corn Producing Counties in the U.S.

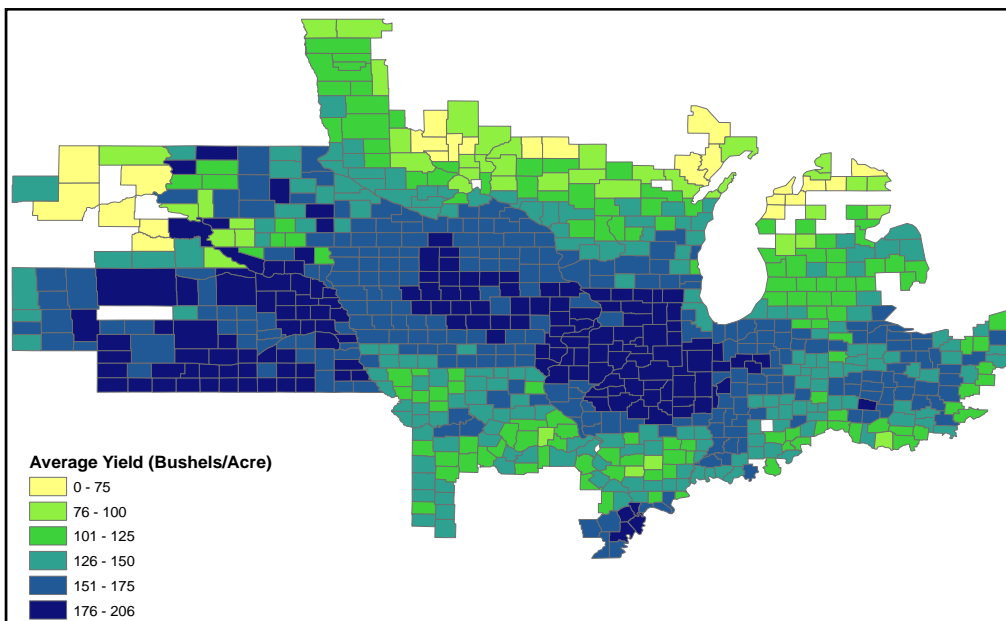


Figure 4.3: Average Yield for Major Corn Producing Counties in the U.S. in 2007

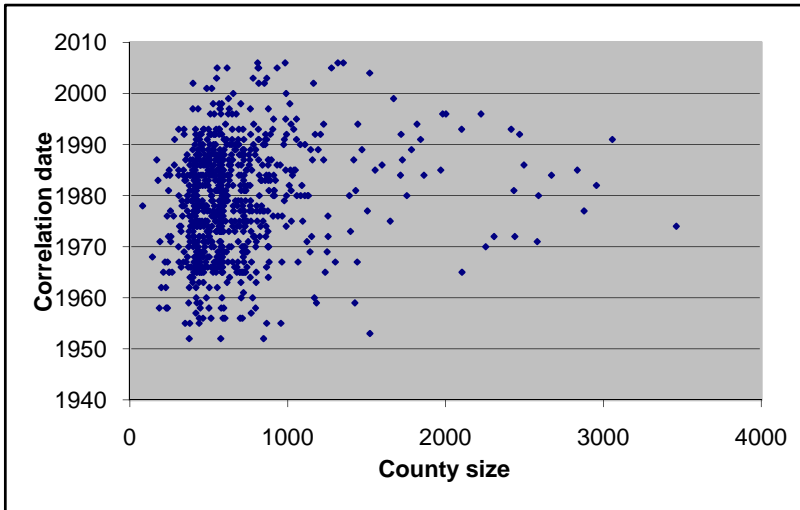


Figure 4.4: Scatter Plot of County Size vs. Correlation Date of Soil Survey

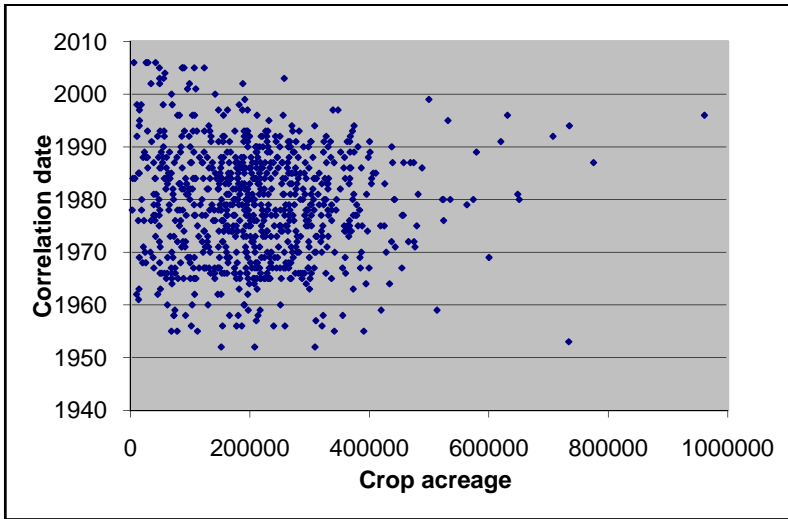


Figure 4.5: Scatter Plot of Crop Acreage vs. Correlation Date of Soil Survey

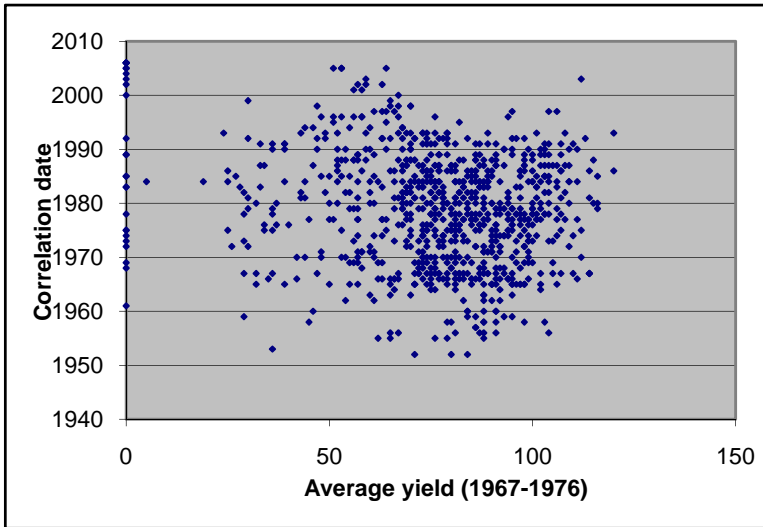


Figure 4.6: Scatter Plot of Average Yield and Dissemination of Soil Survey (Using Correlation Date)

The available data provide a panel data set. The combination of time series with cross-sections enhances the quality and quantity of data in ways that would be impossible using only one of these two dimensions (Gujarati, 2004). Panel data are more informative, provide more variability, have less collinearity among variables, result in more degree of freedom, and give more efficient estimates (Baltagi, 1995). This approach controls for individual unobserved heterogeneity which is not easily detectable in either cross-section or time-series data. A panel data regression is expressed with double subscripts on variables. The model can be represented as:

$$(4.2) \quad Y = \alpha + \beta X'_{it} + u_{it} \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

The subscript i denotes the cross-section dimension and t denotes the time-series dimension. In this analysis model i represents counties and t represents years. The error term in panel data analysis can be decomposed into two components:

$$(4.3) \quad u_{it} = \mu_i + v_{it}$$

where μ_i denotes the unobservable county specific error and ν_{it} denotes idiosyncratic error. The error term μ_i does not change over time and accounts for any county specific effect that is not included in the regression. The error term ν_{it} varies by counties and year.

Generally two types of models are used for panel data analyses: fixed effects models and random effects models. In the fixed effects model, the μ_i are assumed to be fixed parameters to be estimated and the ν_{it} independent and identically distributed $IID(0, \sigma_v^2)$. The fixed effects model consists of too many parameters and suffers from a loss of large degrees of freedom. Loss of degrees of freedom can be avoided if the individual effect, μ_i , can be assumed to be random as in the random effects model. In this case both μ_i and ν_{it} are $IID(0, \sigma_v^2)$ and μ_i are assumed independent of ν_{it} . Also, the independent variables, X_{it} , are independent of μ_i and ν_{it} for all i and t . The random effects model is appropriate when the individuals are selected randomly from a large population (Wooldridge, 2002). Fixed effects models are usually much more convincing than random effects models for policy analysis based on aggregated data (Wooldridge, 2006). The fixed effects model is employed to estimate the regression equation for major counties producing major crops used in this study. The use of a fixed-effects panel estimator allows us to interpret the regression coefficient estimate of an increase in yield for major crops as a measure of soil survey benefits on crop production as soil survey information is made available.

4.2.1. Fixed Effects Method

A fixed effects model allows each county to serve as its own control. This is accomplished by first comparing the variations within counties, and then averaging the differences across all the counties in the sample (Allison, 2005).

There is a trade-off between bias and sampling variability when choosing fixed effects model. A fixed effects model gives less biased estimates at the cost of greater sampling variability. Fixed effects models ignore the between-county variation and deal with only within-county variation. Ignoring the between-county variation can produce higher standard errors than those produced by models using both within and between county variations. Since there is a chance that between-county variation can be influenced by unobserved county characteristics such as policy effects, ignoring the between-county variations may give unbiased estimates.

The data are sorted into a cross section of time-series before analysis using a fixed-effects panel estimator. Thus the data set comprises T observations for each of N counties. Formulation of a fixed effects model assumes that the variation across counties can be captured in the constant term. Each individual county-specific constant is treated as an unknown parameter to be estimated. The equation estimated is:

$$(4.4) \quad y_{it} = \alpha_i + \beta X'_{it} + u_{it} \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

where y_{it} is crop yield in county i in year t , β is a vector of coefficients, X_{it} is a vector of independent variables, and u_{it} is an error term for each county-year observation. The term α_i is a county-specific constant (which is allowed to be unique for each county), α_i represents differences between counties that are stable over time and not accounted for by other variables that do not vary over time such as the productivity index. In a fixed effects

model, α_i terms are considered as fixed parameters, one for each county. In a fixed effects model independent variables, X_{it} , may be correlated with the individual effects, α_i (Hsiao, 2003).

OLS method could be used to estimate the parameters, but applying it with dummy variables for the α_i terms is tedious to compute. However, identical estimates could be obtained by conditioning out the α_i terms and applying the OLS method on deviation scores (Allison, 2005). Conditioned variables are obtained by computing the means for both dependent and independent variables that are varying in each county and for each year.

$$(4.5) \quad \bar{y}_i = \frac{1}{n_i} \sum_t y_{it}$$

$$(4.6) \quad \bar{x}_i = \frac{1}{n_i} \sum_t x_{it}$$

where t is the number of measurements for county i. The county-specific mean is subtracted from the observed values of each variable:

$$(4.7) \quad y_{it}^* = y_{it} - \bar{y}_i$$

$$(4.8) \quad x_{it}^* = x_{it} - \bar{x}_i$$

Using the resulting values, y^* is regressed on x^* .

This model is also called a least squares dummy variable (LSDV) model (Greene, 2003).

The least square estimator of β is given by:

$$(4.9) \quad \hat{\beta} = [X' M_D X]^{-1} [X' M_D y]$$

where $M_D = I - D(D'D)^{-1}D$. X is the entire matrix of independent variables including the county-specific intercepts, y is the vector of observations on county yield, and D is the vector of dummy variables for counties $D = [d_1 \ d_2 \ d_3 \ \dots d_n]$. This equation sums to a least squares regression using the transformed data $X_* = M_D X$ and $y_* = M_D y$. M_D is symmetric, idempotent, and orthogonal to D .

$$M_D = \begin{pmatrix} M^0 & 0 & 0 & \dots & 0 \\ 0 & M^0 & & & \\ 0 & & M^0 & & \\ & & & \dots & \\ 0 & 0 & 0 & \dots & M^0 \end{pmatrix}$$

In this formula, $M^0 = I_T - ii' / T$, where I_T is an identity matrix of rank T , i is a $T \times 1$ vector of ones, and T is the number of periods over which the cross-sections are observed. Thus if there are N counties observed for T years each and k explanatory variables including the constant and the fixed effects, then X is a $TN \times k$ matrix, y is a $TN \times 1$ vector, M^0 is a $T \times T$ matrix, I_n is a $N \times N$ identity matrix, and M_D is a $TN \times TN$ matrix. The matrix M_D controls for correlation across the error terms within counties. The least squares regression of $M_D y$ on $M_D X$ is equivalent to a regression of $[y_{it} - \bar{y}_{it}]$ on $[x_{it} - \bar{x}_{it}]$, where \bar{x}_{it} and \bar{y}_{it} are scalar and $K \times 1$ vector of means of y_{it} and x_{it} over T observations for group i .

The county-specific effects α_i capture all time-invariant characteristics of a location in the above fixed effects model. The use of a fixed effects model avoids the problem of omitted variables, since they are included in the fixed effects (Schlenker and

Roberts, 2006). The regression model is first applied to corn yield and then extended to soybeans, wheat, and cotton.

4.2.2. Data Description

USDA National Agricultural Statistics Service (NASS) county level crop yield data for 1936-2007 were obtained from *Quick Stats: Agricultural Data Base* available from the USDA-NASS web site (<http://www.nass.usda.gov/Quickstats/>). Corn, soybean, and wheat yields are measured in bushels per acre per year whereas cotton yields are measured in pounds per acre per year. While some of the county level crop yield data were available from 1935 or before, data for some counties were not available for the entire period. For example, county level corn yield data for Michigan were first published 1942 and for Kansas in 1958. Likewise, county level cotton and wheat yield data starting from 1935 were only available for some states. Table 4.2 shows USDA-NASS winter wheat yield data status on some of the major winter wheat producing states. Table 4.3 shows USDA-NASS cotton yield data status on some of the major cotton producing states. In general, the data were most complete for corn and with less complete period data for the other crops. All available data were included in the analyses that follow.

Table 4.2: USDA-NASS Winter Wheat Yield Data Status

States	Beginning Year Available
Texas	1968
Washington	1972
Nebraska	1956
Iowa	1972
Arkansas	1961

Table 4.3: USDA-NASS Cotton Yield Data Status

States	Beginning Year Available
Texas	1968
Arkansas	1938
Missouri	1941
Louisiana	1954
Tennessee	1948

County level soil survey completion dates were collected and verified from various sources. Information on county level soil survey publications was obtained from the NASIS (National Soil Information System) database. The NASIS database was compared with the county level soil reports available from Evansdale Library of West Virginia University (WVU) and the Agriculture Science Library of WVU. The WVU Libraries are the Federal Depository for the state of West Virginia, so the WVU Library contains all the soil survey published reports. However, some of the records were missing from WVU Libraries. The publication dates for the remaining records in the National Soil Information System (NASIS) database that were not available from WVU Libraries were compared with the National Agricultural Library (NAL) website records. Records for publication dates were electronically available at the NAL website. At last, some of the records were updated from the NRCS website. Some of the manuscripts of soil survey publications were available at the USDA-NRCS website (http://soils.usda.gov/survey/online_surveys/).

The updated soil survey reports inventory was provided additional information for this research. The dates when soil survey reports were correlated and published provide two measures as to the year when soil information is made available. Dummy variables were created for the soil survey information for each county, with a value 0 prior to the availability of soil survey information and a value of 1 for every year after. The basic assumptions for soil information for a linear model are that this information was used each

year after its provision and that this information provided a constant annual impact on crop production. Table 4.4 below shows the number of soil surveys correlated by time frame for the Corn Belt States.

Table 4.4: Number of Soil Surveys Correlated for the Corn Belt States

States	1954-1960	1961-1970	1971-1980	1981-1990	1991-2000	2001-2007
Illinois	5	15	24	34	22	-
Indiana	3	19	38	32	-	-
Iowa	6	16	40	35	2	-
Michigan	3	17	17	21	17	8
Minnesota	5	15	21	24	16	4
Nebraska	1	28	31	26	4	2
Ohio	3	21	28	29	6	-
South Dakota	3	12	25	20	5	-
Wisconsin	7	19	19	12	6	9

Based on previous studies (Kaufmann and Snell, 1997; Garcia et al., 1987; Schroder et al., 1984), possible weather variables that could be used in corn and soybeans studies include pre-season moisture data and monthly precipitation and temperature for June, July, and August. Thus nine weather variables were initially utilized in the corn and soybeans models: minimum temperature, maximum temperature, and total precipitation for June, July, and August for each county each year.

Gridded climate data provided by the Michigan State Chief Climatologist's Office included longitude, latitude (in hundredths of degrees), and the daily value for the grid point for the lower 48 states for each year for June, July, and August. An inverse distance weighted (IDW) technique is used to interpolate measures for the county centroid from the four closest grid points. A neighborhood about the interpolated point is identified and a weighted average is taken of the observation values within this neighborhood. The weights are a diminishing function of distance. IDW methods are based on the assumption that the

interpolating surface should be influenced most by the nearby points and less by the more distant points. Various options are available for IDW interpolation techniques.

Precipitation records can have a short spatial correlation length scale and large variability, whereas the temperature records have a long spatial correlation scale (Shen et al., 2001).

Thus for interpolating the precipitation data more emphasis is on the nearest points.

Temperature and precipitation data are then recorded for each county centroids from the interpolated surface. Monthly weather values were obtained for each county by averaging the daily values.

Past extreme weather events have caused severe crop damage and consequently caused significant economic losses. Most of those weather events that could affect corn and soybeans production are captured by the above discussed nine variables. The effect of the 1993 Mississippi River Valley floods was not captured by these variables. Flooding in the summer months of 1993 affected 16,000 square miles of farmland in the Midwest damaging over 11 million acres of crops (Rozenzweig, 2001). To reflect the unusual nature of the effect of the 1993 flood event, a dummy variable for 1993 is added to the model.

Time trend variables are included to capture patterns of technological change. Past studies have commonly included time trends as the appropriate proxy to estimate the effect of technology on yield (Garcia et al., 1987; Kaylen et al., 1992; Houck and Gallenger, 1976; Menz and Pardey, 1983; and Buller, 1972). Linear and square polynomial trends are used in the model to disentangle technological effects such as fertilizers, hybrids, and pesticides. Selection of these polynomial trends is based on model performance.

A National Commodity Crop Productivity Index (NCCPI) is included in the corn and soybeans yield model. NCCPI, developed by NRCS, is an interpretation in the National Soil Information System (NASIS). The NCCPI provides a measure of the spatially variable soil productivity for a particular crop across major soils (Dobos, Sinclair, and Hipple, 2008). It is derived to interpret natural relationships of soil, landscape, and climate factors in crop response and is only calculated for non-irrigated commodity crops. The NCCPI model, used to develop the NCCPI index, is based on a relative productivity index or ranking over periods of years. The NCCPI index is between one and zero.

Farm gate prices for crop are available from the USDA-NASS web site (http://www.nass.usda.gov/QuickStats/PullData_US.jsp). Since the average yearly price does not differ significantly, national average prices are used. These prices are adjusted for deflation using the Producer Price Index (PPI). Crop management decisions change in response to price variation which could affect the average yield. Buller (1972) stated that prices could also affect per-acre yield. He added that a high crop price stimulates farmers to improve management by applying more inputs such as fertilizer, better weed control, improved tillage, and others that could result in increased yields. However, crop price increases could encourage farmers in both intensive and extensive ways to increase production. One intensive example, as mentioned by Buller, is by applying more inputs such as fertilizers. One extensive example is by adding more land for crop production. So changes in price could have negative or positive aggregate impacts on average yield.

Even though prices are usually determined by the market, crop prices in the U.S. are influenced by government agriculture policies. Government policies on price support

and commodity storage have affected market prices of some crops such as corn and wheat, at times quite significantly. The government, through income support policies, subsidizes farm income by artificially increasing commodity prices. The new target price, facilitated by subsidies, induces farmers to increase their production and their profit. For example, the production of many major crops in the U.S. were restricted by government acreage set-aside requirements until the 1996 Freedom to Farm Act (Just, Hueth, and Schmitz, 2004). The government offered price support payments or higher-than market target prices, to attract farmers holding portion of unused land or land occupied to a particular crop.

4.2.2.1. Other Desired Data

Technology variables based on previous studies (Griliches, 1957; Kaufmann and Snell, 1997; Schroder et al., 1984) that could be employed in this study include hybrid introduction and fertilizer use. However, the information to develop comparable data for these variables across the spatial and temporal dimensions of this study is not available. In the case of hybrids, there are no studies that imply a spatial variation in hybrid introduction. After more than three months attempting to collect fertilizer information, the lack of consistent information became apparent. The sources and reporting basis for fertilizer data vary significantly across both space and time. State level fertilizer information has only been reported by the USDA since 1966.

Alexander and Smith (1990) estimated county level nitrogen and phosphorous fertilizer use for 1945-1985 by disaggregating state-level USDA data (1966, 1976, 1977-1985) to county level. However, they noted that county level estimates of fertilizer use prior to the 1970s should be used with caution. Likewise, the USGS Water Resources Division provided fertilizer sales data for 1986-1991. Since this study employs data for

1936-2007, it is impossible to acquire fertilizer data from the beginning of this period. Based on this approach, consistent time trends are of utmost importance. Projecting fertilizer use data for the earlier period (before 1966) may impose additional errors and bias the results. While inclusion of fertilizer data is preferred, because of the statistical issues, fertilizer data are not included in the analysis for this research.

4.2.3. Crop Yield Model

The empirical model presented is first applied to corn yield response functions. Then the model is extended to soybeans, winter and spring wheat, and upland cotton. Based on the data availability, different mathematical crop formulations are developed for corn, soybeans, wheat, and cotton to estimate the effect of the provision of soil information in crop yield. The mathematical forms of corn yield response functions are:

$$(4.10) \quad \begin{aligned} \text{corn_yield} = & \text{dd_pub Lag_CornPrice year1993 trend1 trend2 NCPPI} \\ & \text{june_ppt july_ppt aug_ppt june_mxt july_mxt aug_mxt} \\ & \text{june_mnt july_mnt aug_mnt} \end{aligned}$$

$$(4.11) \quad \begin{aligned} \text{corn_yield} = & \text{dd_corr Lag_CornPrice year1993 trend1 trend2 NCPPI} \\ & \text{june_ppt july_ppt aug_ppt june_mxt july_mxt aug_mxt} \\ & \text{june_mnt july_mnt aug_mnt} \end{aligned}$$

Soybeans yield response functions are developed and estimated using the same variables produced for corn yield response functions.

$$(4.12) \quad \begin{aligned} \text{soy_yield} = & \text{dd_pub Lag_SoyPrice year1993 trend1 trend2 NCPPI} \\ & \text{june_ppt july_ppt aug_ppt june_mxt july_mxt aug_mxt} \\ & \text{june_mnt july_mnt aug_mnt} \end{aligned}$$

$$(4.13) \quad \begin{aligned} \text{soy_yield} = & \text{dd_corr Lag_SoyPrice year1993 trend1 trend2 NCPPI} \\ & \text{june_ppt july_ppt aug_ppt june_mxt july_mxt aug_mxt} \\ & \text{june_mnt july_mnt aug_mnt} \end{aligned}$$

Wheat yield response functions and cotton yield response functions are developed using only soil survey and trend variables because of the unavailability of other variables³ and the model's initial performance.

$$(4.14) \quad \text{yield} = \text{dd_pub trend1}$$

$$(4.15) \quad \text{yield} = \text{dd_pub trend1 trend2}$$

$$(4.16) \quad \text{yield} = \text{dd_pub lag_price trend1 trend2}$$

$$(4.17) \quad \text{yield} = \text{dd_corr trend1}$$

$$(4.18) \quad \text{yield} = \text{dd_corr trend1 trend2}$$

$$(4.19) \quad \text{yield} = \text{dd_corr lag_price trend1 trend2}$$

Table 4.5 summarizes the variables employed in corn, soybeans, wheat, and cotton yield response model.

³ Weather data were not available for wheat and cotton production region.

Table 4.5: Definition and Data Sources for Variables Used in the Crop Models

Variable	Definition	Source of Data
dd_pub	Dummy variable for soil survey publication date	NASIS, WVU Library and NAL
dd_corr	Dummy variable for soil survey correlation date	NASIS, WVU Library and NAL
Lag_CornPrice	Lag Corn price	USDA- NASS
Lag_SoyPrice	Lag Soy price	USDA- NASS
Lag_Price	Lag price for wheat or cotton	USDA- NASS
year1993	Dummy variable for year 1993	
<i>Trend variables</i>		
trend1	Linear time trend	
trend2	Quadratic time trend	
<i>Weather Variables</i>		
june_ppt	June precipitation	Michigan State Chief Climatologist's Office
july_ppt	July precipitation	Michigan State Chief Climatologist's Office
aug_ppt	August precipitation	Michigan State Chief Climatologist's Office
june_mnt	June minimum temperature	Michigan State Chief Climatologist's Office
july_mnt	July minimum temperature	Michigan State Chief Climatologist's Office
aug_mnt	August minimum temperature	Michigan State Chief Climatologist's Office
june_mxt	June maximum temperature	Michigan State Chief Climatologist's Office
july_mxt	July maximum temperature	Michigan State Chief Climatologist's Office
aug_mxt	August maximum temperature	Michigan State Chief Climatologist's Office

The major crop producing states were selected based on the USDA-NASS reports “Crop Production 2007 Summary” and “Crop Production 1996 Summary” (USDA-NASS, 1997 and USDA-NASS, 2008). Table 4.6 - 4.10 present lists of the top ten crop producing states for corn, soybeans, wheat, and cotton for the U.S. in 2007⁴. Since irrigated crop yield per acre is primarily influenced by irrigation, only non-irrigated crop production was considered in this analysis. Non-irrigated yield per acre per crop is influenced by natural

⁴The states in these tables may not be included in the dataset for the analysis. These tables represent the total crop production in 2007. However, only non-irrigated portion and counties having at least 20 years observations were included in the analysis. The reasons why these states are not included in the analysis are given in Appendix B.

factors such as weather and soil, level of technology, and management practices (Buller, 1972).

Most of the counties from the top ten corn producing states are included in the corn yield model for 1936-2007. However, for some counties USDA-NASS yield data were not available in 1935. For example, county level corn yield data for Michigan are available starting in 1942 and for Kansas starting in 1958. Thus the data set provides a unbalanced panel. Figure 4.7 shows the major corn producing counties from ten states included in this study.

Similarly, most of the counties are included in the soybean model for the same states as used in the corn model. This area represents more than 80% of U.S. soybean production. Including counties with few observations in the analysis may lead to biased estimates. Thus only counties having at least 20 years of observations were included in the soybean model. Figure 4.8 shows the major soybean producing counties from ten states included in this study.

Table 4.6: Top Ten U.S. Corn Producing States in 2007

State	Bushels (in thousands)	Percent of U.S.
Iowa	2,368,350	18.12%
Illinois	2,283,750	17.47%
Nebraska	1,472,000	11.26%
Minnesota	1,138,800	8.71%
Indiana	987,350	7.55%
South Dakota	544,500	4.16%
Ohio	541,500	4.14%
Kansas	518,000	3.96%
Missouri	461,500	3.53%
Wisconsin	442,800	3.39%
		Total - 82.29%
U.S. Total	13,073,893,000	

(Source: USDA-NASS, 2008. "Crop Production 2007 Summary" Available on line at <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-11-2008.pdf>)

Table 4.7: Top ten U.S. Soybeans Producing States in 2007

State	Bushels (in thousands)	Percent of U.S.
Iowa	438,780	16.97%
Illinois	350,450	13.56%
Minnesota	252,150	9.75%
Indiana	210,600	8.15%
Ohio	194,110	7.51%
Nebraska	190,385	7.36%
Missouri	168,350	6.51%
South Dakota	133,560	5.17%
North Dakota	104,650	4.05%
Arkansas	100,440	3.89%
		Total - 82.91%
U.S. Total	2,585,207	

(Source: USDA-NASS, 2008. "Crop Production 2007 Summary" Available on line at <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-11-2008.pdf>)

Table 4.8: Top ten U.S. Winter Wheat Producing States in 2007

State	Bushels (in thousands)	Percent of U.S.
Kansas	283,800	18.72%
Texas	140,600	9.27%
Washington	108,160	7.13%
Oklahoma	98,000	6.46%
South Dakota	95,040	6.27%
Colorado	94,000	6.20%
Nebraska	84,280	5.56%
Montana	83,220	5.49%
Idaho	51,830	3.42%
Illinois	50,730	3.35%
		Total - 71.88%
U.S. Total	1,515,989	

(Source: USDA-NASS, 2008. "Crop Production 2007 Summary" Available on line at <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-11-2008.pdf>)

Table 4.9: Top ten U.S. Spring Wheat Producing States in 2007

State	Bushels (in thousands)	Percent of U.S.
North Dakota	234,000	48.85%
Minnesota	77,550	16.19%
Montana	55,200	11.52%
South Dakota	52,260	10.91%
Idaho	30,600	6.39%
Washington	20,562	4.29%
Oregon	6,360	1.33%
Colorado	1,520	0.32%
Utah	420	0.09%
Wisconsin	280	0.06%
		Total - 99.94%
U.S. Total	479,047	

(Source: USDA-NASS, 2008. "Crop Production 2007 Summary" Available on line at <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-11-2008.pdf>)

Table 4.10: Top ten U.S. Upland Cotton Producing States in 2007

State	Bales (in thousands)	Percent of U.S.
Texas	8,100	44.49%
Arkansas	1,880	10.33%
Georgia	1,650	9.06%
Mississippi	1,330	7.30%
North Carolina	785	4.31%
Missouri	770	4.23%
Louisiana	690	3.79%
California	630	3.46%
Tennessee	615	3.38%
Arizona	500	2.75%
		Total - 90.34%
U.S. Total	18,208	

(Source: USDA-NASS, 2008. "Crop Production 2007 Summary" Available on line at <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-11-2008.pdf>)

Since yield data for most of the wheat producing counties were not available over the entire time period, only counties having at least 20 years of observations were included in both of the wheat models. A total of 199 counties from top spring wheat producing states were included in the spring wheat model and 486 counties from the top winter wheat producing states are included in the winter wheat model (Figure 4.9 and 4.10).

Similarly, 190 counties from the top cotton producing counties are included in the upland cotton model (Figure 4.11), and the model includes only the counties with at least 20 years observations. A table and graph showing the available number of counties for each year for the counties selected for this research is included as Appendix A.

Table 4.11 - 4.13 summarize the descriptive statistics of the variables in the corn, soybeans, wheat, and cotton models based on the number of observations used in the corn, soybeans, wheat, and cotton models, respectively.

Table 4.11: Descriptive Statistics - Corn Yield Model

Variable	Mean	Standard Deviation	Minimum	Maximum
Yield	77.79	40.72	1.00	204.00
NCCPI	0.50	0.17	0	0.88
Lag_Cornprice	2.78	0.67	1.61	5.05
<i>Weather Variables</i>				
june_ppt	1.36	0.58	0.080	4.97
july_ppt	1.15	0.55	0.06	5.75
aug_ppt	1.10	0.53	0.03	5.13
june_mnt	57.31	4.22	40.94	70.44
july_mnt	61.84	3.73	48.39	72.77
aug_mnt	60.05	3.96	44.74	71.72
june_mxt	80.68	4.13	64.79	97.08
july_mxt	85.49	4.22	69.05	104.23
aug_mxt	83.70	4.45	69.14	100.97

Table 4.12: Descriptive Statistics – Soybean Yield Model

Variable	Mean	Standard Deviation	Minimum	Maximum
SoyYield	28.54	10.59	2.00	64.00
NCCPI	0.54	0.15	0	0.88
Lag_SoyPrice	5.93	1.04	3.61	8.64
<i>Weather Variables</i>				
june_ppt	1.38	0.58	0.10	4.97
july_ppt	1.20	0.55	0.07	5.75
aug_ppt	1.14	0.53	0.03	5.13
june_mnt	57.95	3.78	42.44	70.44
july_mnt	62.22	3.39	49.08	72.77
aug_mnt	60.35	3.68	45.72	71.72
june_mxt	80.93	3.86	65.67	97.08
july_mxt	85.22	3.80	70.71	101.55
aug_mxt	83.43	4.08	70.08	98.53

Table 4.13: Descriptive Statistics – Wheat and Cotton Yield Model

Variable	Mean	Standard Deviation	Minimum	Maximum
<i>Winter Wheat Model</i>				
Yield	29.46	11.48	2.5	112.50
Lag_price	3.42	0.76	2.33	5.12
<i>Spring Wheat Model</i>				
Yield	24.73	10.65	1.00	105.00
Lag_price	3.74	0.81	2.30	5.89
<i>Upland Cotton Model</i>				
Yield	452.16	216.25	34.00	1206.00
Lag_price	83.43	4.08	70.08	98.53

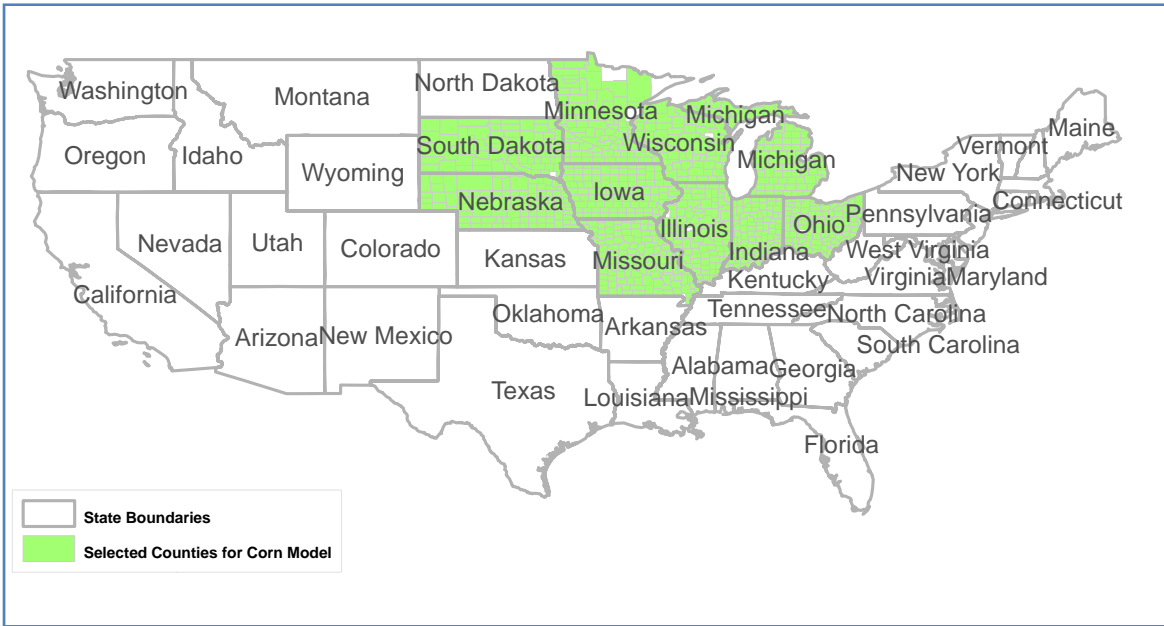


Figure 4.7: Selected Corn Producing Counties

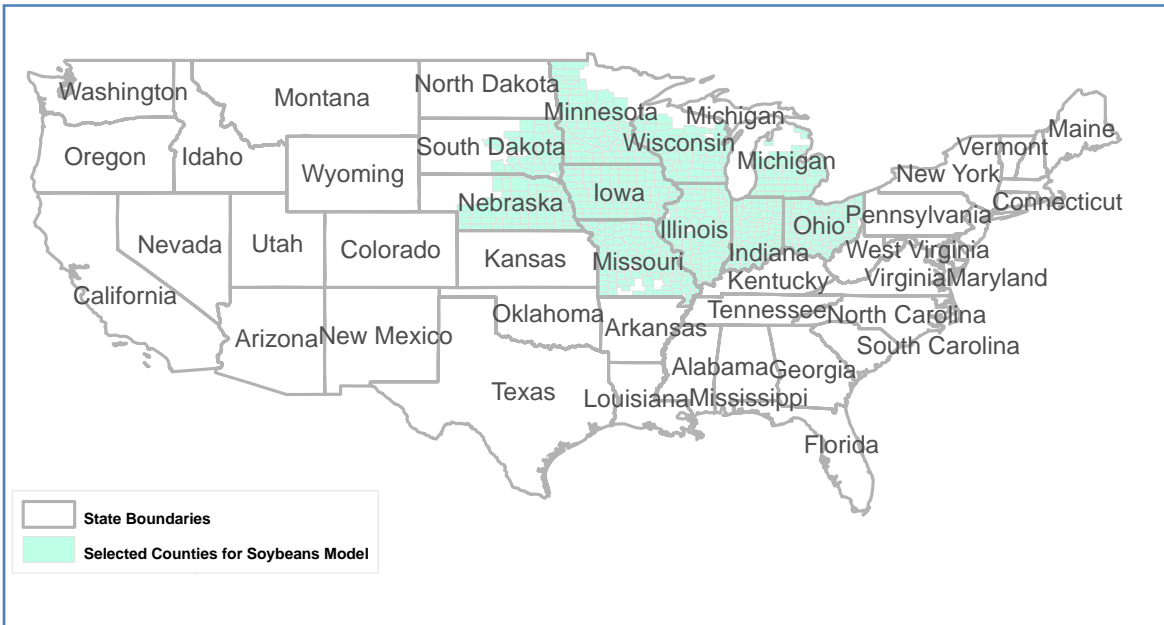


Figure 4.8: Selected Soybean Producing Counties

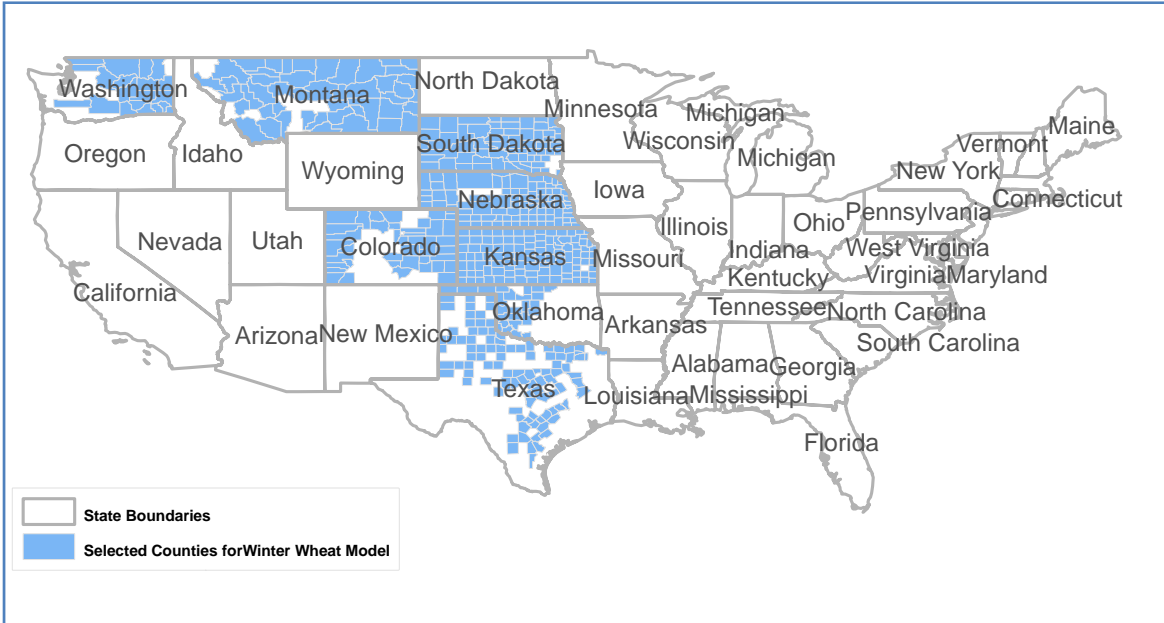


Figure 4.9: Selected Winter Wheat Producing Counties

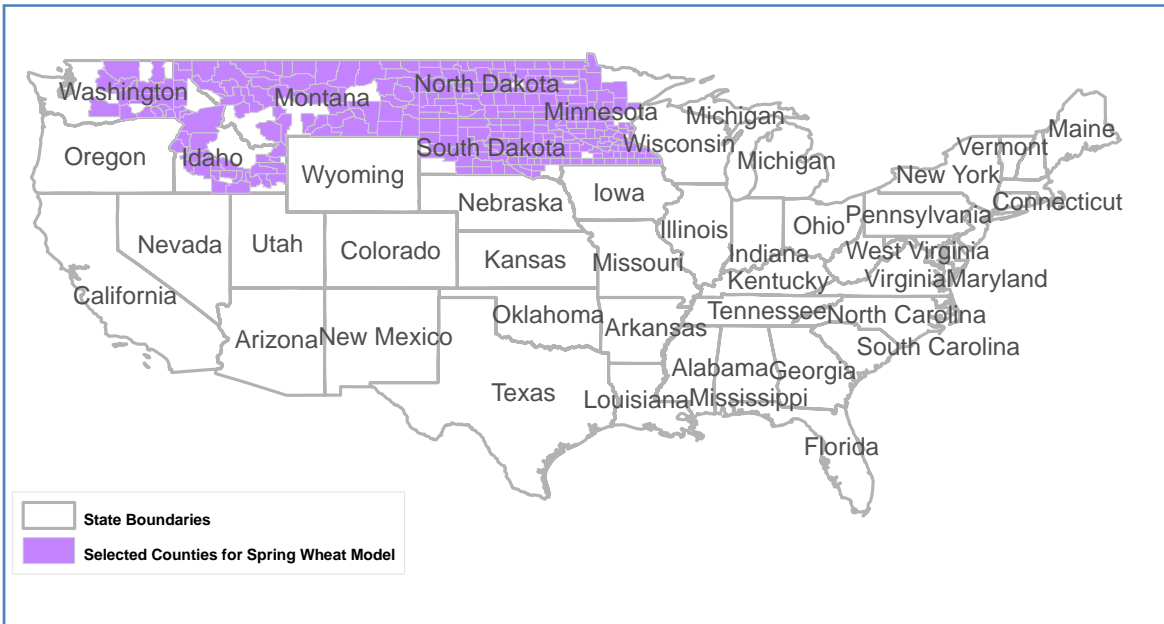


Figure 4.10: Selected Spring Wheat Producing Counties

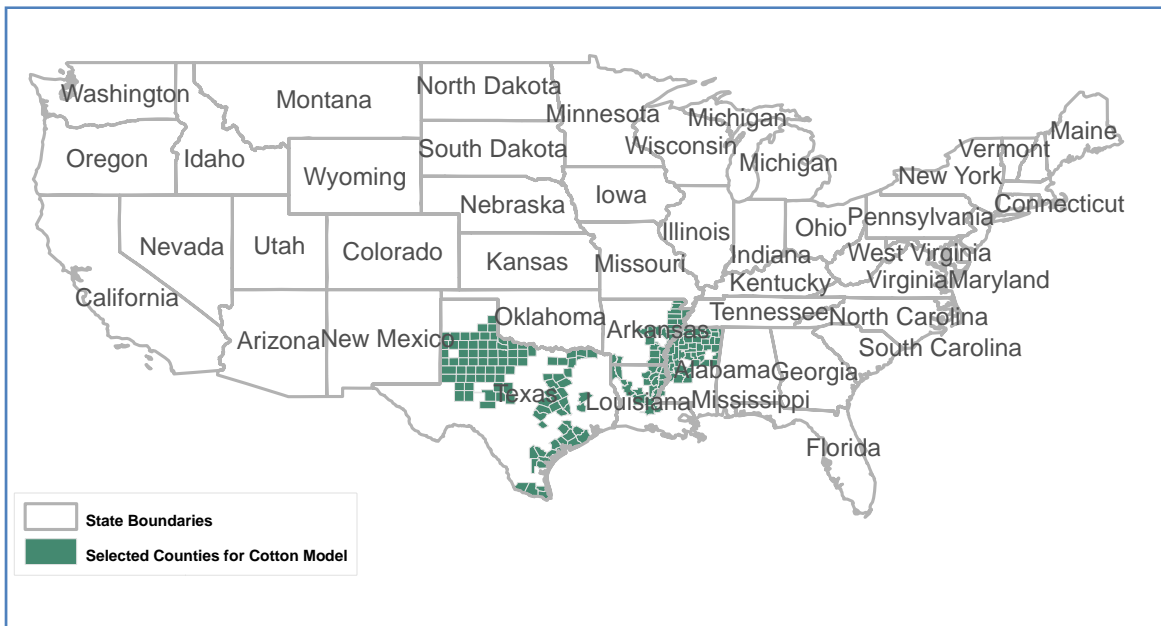


Figure 4.11: Selected Cotton Producing Counties

4.3. BENEFIT-COST ANALYSIS

The previous section provided a detailed methodological approach for soil survey benefit estimation in selected crop yields. This section utilizes the estimates developed in the previous section to estimate the net partial benefit of the soil survey program.

A benefit-cost analysis (BCA) is an economic tool commonly used to analyze public policies and regulatory decisions. BCA is used to estimate the net economic value of a given project or policy and thus can be very useful in gauging the effectiveness of any government program. It converts all the benefits and cost into a single monetary metric to evaluate all the benefits and costs of the project. If the benefits outweigh the costs, the project improves economic efficiency. If the costs outweigh the benefits, the project decreases economic efficiency. The economic viability of the NCSS program can be evaluated through a BCA. BCA can also be used to estimate the partial benefits of the soil survey program.

Benefits derived from the soil survey program can be categorized as past, present, or future. Past benefits are the benefits achieved in the past from past investments. Only past benefits of the NCCC program for corn and soybean production are used in BCA. The economic benefit of the NCSS program for crop production is estimated by changing the values of increased amount of yield attributable to the availability of soil information to constant dollar values. Benefits are estimated at the county level and aggregated for the study region. However, costs are estimated at the state level. The value of increased yields less estimated production costs provides a lower-bound estimate of the economic benefits of the NCSS.

4.3.1. Methodological Approach

In order to evaluate the economic viability of the NCSS program, two separate benefit-cost analyses are employed. The first BCA is for the scenario where benefits are assumed to be accrued after the soil survey is correlated. The term correlation here refers to the final correlation date marks, the date when the entire survey has been mapped, verified, and potentially would be available in the form of an interim report. In other words, soil data is available but not in a formal polished publication. The second BCA is for the scenario where benefits are assumed to be accrued only after the publication of soil survey reports. This could be as many as 1 to 15 years after the final correlation date. This is the date that the hard copy publication would be available to the public.

Benefits are estimated at the county level, and costs are estimated at the state level. Aggregate inter-temporal values across the 1950-2007 were considered in the analysis. Since it takes three to six years to complete a soil survey, benefits were estimated only after 1954 for publication date scenario, and after 1952 for correlation date scenario and

costs were estimated for 1950-2007. The benefit-cost analysis is based on the various assumptions illustrated in Table 4.14.

Sensitivity analyses are conducted for both scenarios using various discount rates to evaluate the impact of the NCSS program. The choices of discount rates are further detailed in the following section.

Table 4.14: Basic Assumptions for Benefit-Cost Analysis

Assumptions

- All the farmers use soil survey information provided by the NCSS program.
 - For the correlation date scenario, benefits accrue immediately after the soil survey is correlated and approved.
 - For the publication date scenario, benefits accrue immediately after the publication of the soil survey report.
 - Benefits and costs have accrued up to 2007.
 - Yearly total cost of soil survey is equivalent to federal budget allocation for soil survey program.
 - Each state receives the same portion of the total U.S. federal budget allocation for the NCSS program, based on years 1987-1989.
-

4.3.2. Choice of Discount Rate

When the benefits and costs of a project accrue over a period of years, all the monetary values are discounted to a single point of time. According to Lang and Marino, discounted cash flow calculates the value today of a cash sum to be realized in the future (Lang and Marino, 1993). Discounting future benefits or costs over time reflects society's time preference for money and the discount rate reflects society's present preference of consumption relative to future consumption (Freeman, 2003). The discount rate reflects the opportunity cost of capital, valued in terms of investment, which accounts for the

alternative use of the capital (Berlage and Renard, 1985). Even though discounting is a necessary part of benefit-cost analysis, application of the discount rate creates ambiguities for benefit-cost analysis. The choice of discount rate has a significant effect on the evaluation of benefits and costs when the time horizon is long. A high discount rate lowers the relative value of benefits received in the future. Just, Hueth, and Schmitz (2004) suggest a range of 2- 4% is an appropriate social discount rate. The U.S. Office of Management and Budget (1992) states that benefit-cost analyses of proposed investments and regulations should report net present value using a real discount rate of 7%. Therefore, discount rates of 2%, 4%, and 7% are used for the public benefit-cost analysis.

4.3.3. Evaluation Technique

The historical net-benefit of soil survey information in agricultural production is calculated in this research. Costs are estimated from available state level soil survey historical budget allocation and yearly total U.S. budgets. Benefits, however, are much more complicated to quantify. Benefits are based on productivity gains of major crops related to the provision of soil information at the county level. The value of increased crop yields less estimated increased productions costs provides a lower-bound estimate of the economic benefits of the NCSS program.

Net present value (NPV) and benefit-cost ratio are utilized to evaluate the effectiveness of the NCSS program. Benefits are expected to be positive for economically viable projects. Microsoft Excel is used for the analysis. Mathematically, the formula for calculating the NPV of the past investment is similar to calculating the future value of present value:

$$(4.20) \quad NPV = \sum_i \sum_{t_i=ssd_i}^{2007} (B_t * (1+r)^{(2007-t)} * (1+f)^{(2007-t)}) - \sum_{t=1950}^{2007} C_t * (1+r)^{(2007-t)} * (1+f)^{(2007-t)}$$

where, B = sum of all benefits,

C = sum of all costs,

ssd_i = year when soil survey information was available

for county i ,

r = discount rate,

f = inflation rate, and

t = time period.

The benefit for provision of soil survey information is estimated by multiplying the increased yield in crop yield attributed to soil survey information by crop price.

$$(4.21) \quad B_t = \sum_j (P * \delta * L)$$

So, NPV can be expressed by,

$$(4.22) \quad NPV = \sum_i \sum_{t_i=ssd_i}^{2007} \sum_j ((P * \delta * L) * (1+r)^{(2007-t)} * (1+f)^{(2007-t)}) - \sum_{t=1950}^{2007} C_t * (1+r)^{(2007-t)} * (1+f)^{(2007-t)}$$

where P = price of crop per bushel,

δ = increase in crop yield attributed to soil information in
bushels per acre area per year,

L = acres of land harvested, and

C = total state level budget allocation for soil survey program.

Subscript i represents the county and j represents the crop.

The benefits and costs of the NCSS program are presented in constant dollar terms. The Gross Domestic Product (GDP) inflator is employed to adjust the effects of inflation on the cash flow analysis. Since it is ambiguous whether producers or consumers benefit from the increased crop production and the cost estimate for this analysis is an aggregate cost, the GDP inflator seems to be more appropriate than other available price index

inflaters. The GDP inflator measures the price of all the goods and services included in GDP. Unlike the Consumer Price Index (CPI) derived only from a representative consumer's basket of goods and the Producer Price Index (PPI) derived only from the producer's perspective, the GDP inflator is derived from an array of the entire collection of goods and services. GDP information is available from the Bureau of Economic Analysis, U.S. Department of Commerce (<http://www.bea.gov/national/#gdp>). Currently the Bureau of Economic Analysis uses the year 2000 as the base year.

4.3.4. Benefit Estimation

The benefits in the analyses are derived from the increase in crop yield that is attributable to the availability of soil information. Benefits are based on productivity gains of major crops related to the provision of soil information at the county level. The productivity gains are estimated using a fixed effects panel data approach to estimate the increase in county level crop yields attributable to the availability of improved soil information. The benefits of the NCSS program to increased agriculture productivity are estimated by aggregating the benefits to major crops that have accrued from the provision of soil information.

The monetary value of soil survey information benefits are estimated by multiplying the increased yield in crop yield attributed to soil survey information by crop price and number of acres of land harvested for that particular crop. For example, if soil survey information for a county was available after 1980, then the benefit is estimated by aggregating the inter-temporal benefits. The number of acres of land harvested each year after 1980 is multiplied by the estimated increased yield and crop price of that year to

estimate the value of soil survey for that county. Therefore, benefits are the sum of all county level inter-temporal benefits of increased yield for crops.

4.3.5. Cost Estimation of Soil Information

The major cost in soil survey is the staff salaries and related overhead which includes accommodation and fieldwork allowances (Dent and Young, 1981). Other costs include equipment costs, laboratory costs, and publication costs. All these cost are related to the amount of field work and scale of mapping. Even though, the benefits are estimated at the county level, historical budget allocations are not documented by the county level . Despite the initial impression that costs (based on state level agency budget information) are relatively straight forward and historical budget allocations would be available from the NRCS, it was not possible to acquire all the state level historical budget allocations for the soil survey program.

State level budget allocations for 1987-1989 were obtained by personal contact (Paul Benedict, Soil Scientist and Program Manager of Soil Survey Division, NRCS) and for 2005-2007 were obtained from the NRCS website (<http://www.nrcs.usda.gov/programs/>). Yearly total soil survey budget allocations for the U.S. were available from the yearly volume of Budget of the U.S. Government (U.S. Bureau of the Budget, 1950-2009). For other remaining years state level budget allocations are approximated based on the state level budget allocations for 1987-1989 and the yearly total soil survey budget allocations for the U. S. Since costs of the soil survey program for 1950-2007 are considered in this analysis, only state level budget allocations for 1987-1989 are used as an approximation to avoid potential bias due to the greater weight of the more recent budget allocation.

Soil survey cost information for other contributors such as the state, county, and private parties were not available. Thus yearly soil survey cost estimates used in this study only consider the total federal budget allocation to the soil survey program and ignore other contributors. Thus the cost estimates of soil survey used in this study underestimates the real cost of the soil survey program.

For the analysis, estimates of state level budget allocation for the soil survey program are used to approximate the total cost of soil information in ten states (Table 4.15). Discount rates of 2%, 4%, and 7% are used to estimate the present value.

Table 4.15: Estimates of State Level Federal Budget Allocation of the NCSS Program

States	Total Budget for 1950-2007 (in million dollar)		
	2%	4%	7%
Illinois	\$254	\$489	\$1,467
Indiana	\$140	\$269	\$808
Iowa	\$206	\$397	\$1,192
Michigan	\$209	\$402	\$1,209
Minnesota	\$313	\$603	\$1,809
Missouri	\$300	\$578	\$1,734
Nebraska	\$152	\$293	\$881
Ohio	\$147	\$283	\$850
South Dakota	\$153	\$294	\$884
Wisconsin	\$153	\$295	\$885
Total Budget	\$2,030	\$3,907	\$11,724

CHAPTER 5

EMPIRICAL RESULTS AND ANALYSIS

5.1. EMPIRICAL ESTIMATION FROM CROP MODELS

Estimation results for the crop models presented in previous chapters and the analysis of those results are presented in this chapter. The crop models (equation 4.10 through 4.19) provide the structure for the relationship between the endogenous variable, crop yield, and exogenous variables including soil survey information, lag price, productivity index, weather variables, and time trend. The equations are estimated using the fixed effects panel data method using the SAS software package. The total number of observations, number of years, and number of counties included in each model are summarized in Table 5.1.

Table 5.1: Number of Observations Included in Each Model

Model	Number of Observations	Number of Years	Number of Counties
Corn (Correlation Date)	61182	72 (1936-2007)	885
Corn (Publication Date)	61017	72 (1936-2007)	880
Soybeans (Correlation Date)	49576	72 (1936-2007)	787
Soybeans (Publication Date)	49594	72 (1936-2007)	787
Winter Wheat (Correlation Date)	19489	72 (1936-2007)	486
Winter Wheat (Publication Date)	19303	72 (1936-2007)	470
Spring Wheat (Correlation Date)	13537	72 (1936-2007)	276
Spring Wheat (Publication Date)	13950	72 (1936-2007)	279
Cotton (Correlation Date)	6009	36 (1972- 2007)	190
Cotton (Publication Date)	5841	36 (1972- 2007)	184

5.1.1. *Corn Yield Model*

The corn yield equations were estimated as a function of a soil information variable representing the time that soil survey information became available to decision makers, a productivity index, the lagged price, weather variables, and time trend variables. The coefficients for the dummy variables that indicated the availability of soil survey information were found to be statistically significant for both of the corn yield models (equation 4.10 and 4.11 using the publication date and correlation date, respectively) and the estimates obtained were consistent. The results are reported in Table 5.2 and Table 5.3. Adding or removing other variables such as linear and higher order time trend, NCCPI, or weather variables from both of the models did not significantly change the coefficient that captured the effect of the availability of soil survey information. Results for the reduced models are included as Appendix C.

Coefficients for the provision of soil survey information were positive and highly significant. Since the productivity index is county specific and time invariant, it did not change the result of the regression analysis. Since the effects of the NCCPI were not estimable using a fixed effects model, the NCCPI is not included in the final model. Results including the NCCPI are included in Appendix D. The time trend variables and all weather variables except June maximum temperature are statistically significant in both of the corn yield models. The maximum temperature variables had a negative sign as expected which indicates that a high maximum temperatures adversely affects corn yield.

The lag price coefficient was positive, indicating that an increase in price leads to an increase in corn yield in the following year. Corn model results support price effects on the intensive margin. Farmers respond to price increases by changing their management activities to increase yield. The dummy variable coefficient for the 1993 Mississippi flood

had a negative sign and was highly significant which captures the serious impacts of the flood in 1993 and improve the overall model results.

Table 5.2: Corn Yield Model Results (Publication Date)

Parameter	Estimate	Standard Error	p-value
dd_pub	1.812	0.229	<.0001
Lag_Cornprice	0.606	0.134	<.0001
year1993	-30.592	0.538	<.0001
trend1	0.839	0.014	<.0001
trend2	0.009	0.000	<.0001
june_ppt	-0.369	0.141	0.009
july_ppt	5.547	0.161	<.0001
aug_ppt	2.024	0.158	<.0001
june_mxt	-0.012	0.042	0.773
july_mxt	-1.413	0.049	<.0001
aug_mxt	-0.895	0.046	<.0001
june_mnt	0.184	0.047	<.0001
july_mnt	1.115	0.056	<.0001
aug_mnt	-0.448	0.050	<.0001
R-square	0.871	Root MSE	14.679
Coeff Var	18.659		

Table 5.3: Corn Yield Model Results (Correlation Date)

Parameter	Estimate	Standard Error	p-value
dd_corr	2.016	0.226	<.0001
Lag_Cornprice	0.583	0.134	<.0001
year1993	-30.794	0.538	<.0001
trend1	0.826	0.014	<.0001
trend2	0.009	0.000	<.0001
june_ppt	-0.368	0.141	0.009
july_ppt	5.545	0.161	<.0001
aug_ppt	2.031	0.158	<.0001
june_mxt	-0.020	0.042	0.642
july_mxt	-1.417	0.048	<.0001
aug_mxt	-0.897	0.046	<.0001
june_mnt	0.189	0.047	<.0001
july_mnt	1.123	0.056	<.0001
aug_mnt	-0.437	0.050	<.0001
R-square	0.870	Root MSE	14.678
Coeff Var	18.699		

5.1.2. Soybean Yield Model

The soybean yield equations were estimated using the same structure as the corn yield equations. The results are reported in Table 5.4 and Table 5.5. As in the corn models, the coefficients for the dummy variables that captured the availability of soil survey information in both the soybeans models (equation 4.12 and 4.13 using the publication date and correlation date, respectively) were found statistically significant and the estimates obtained were similar. Results from both of the models using publication date or correlation date were similar and consistent. Adding or removing other variables such as linear and higher order time trend, NCCPI, or weather variables from both of the models did not significantly change the coefficient for soil survey. Results for the reduced models are included in Appendix C.

Coefficients for the provision of soil survey information were positive and significant. Like the corn yield model, the NCCPI variable did not change the result of the

regression analysis. As noted, the effects of the NCCPI variable were not uniquely estimable using a fixed effects model, so NCCPI was not included in the final model. Time trend and all the weather variables were statistically significant in soybeans model. July maximum temperature and August maximum temperature variables had negative signs, which means high temperatures in July and August adversely affect soybean yield.

The coefficients for lag price for soybeans model had a negative sign. This suggests that unlike in the case of corn production, increasing own price affects the extensive margin. Farmers may add marginal land for soybean production, resulting in a decrease in overall average county yield. Dummy variable coefficients for the 1993 Mississippi flood had negative signs and were highly significant. However, the effect was approximately six times higher for corn yield than soybeans yield, most likely due to the additional effects in the early season on corn.

Table 5.4: Soybean Yield Model Results (Publication Date)

Parameter	Estimate	Standard Error	p-value
dd_pub	0.337	0.072	<.0001
Lag_Soyprice	-0.511	0.046	<.0001
year1993	-5.222	0.166	<.0001
trend1	0.315	0.005	<.0001
trend2	0.001	0.000	<.0001
june_ppt	-0.110	0.045	0.014
july_ppt	0.569	0.050	<.0001
aug_ppt	1.601	0.049	<.0001
june_mxt	0.066	0.015	<.0001
july_mxt	-0.511	0.016	<.0001
aug_mxt	-0.548	0.016	<.0001
june_mnt	0.174	0.016	<.0001
july_mnt	0.615	0.018	<.0001
aug_mnt	0.276	0.016	<.0001
R-square	0.848	Root MSE	4.215
Coeff Var	14.659		

Table 5.5: Soybean Yield Model Results (Correlation Date)

Parameter	Estimate	Standard Error	p-value
dd_pub	0.212	0.071	0.003
Lag_Soyprice	-0.236	0.021	<.0001
year1993	-5.246	0.166	<.0001
trend1	0.346	0.005	<.0001
trend2	0.000	0.000	<.0001
june_ppt	-0.100	0.045	0.025
july_ppt	0.580	0.050	<.0001
aug_ppt	1.624	0.050	<.0001
june_mxt	0.079	0.015	<.0001
july_mxt	-0.515	0.016	<.0001
aug_mxt	-0.565	0.016	<.0001
june_mnt	0.162	0.016	<.0001
july_mnt	0.626	0.018	<.0001
aug_mnt	0.288	0.016	<.0001
R-square	0.844	Root MSE	4.219
Coeff Var	14.728		

5.1.3. Wheat Yield Model

The results for wheat yield models were not as promising as for corn and soybeans (Tables 5.6 through 5.9). Results for the wheat models using equation 4.14, with a dummy variable for publication date and linear time trend, had a positive but not statistically significant coefficient for the availability of soil information for the winter wheat model but negative and statistically significant for spring wheat. The results for equation 4.15, a yield model with dummy variables for publication date and linear and quadratic time trend, gave a coefficient for the soil information variable that was positive for winter wheat and negative for spring wheat. Neither coefficient was statistically significant. The results for equation 4.16, a yield model with dummy variables for publication date, lag own price, linear and quadratic time trend, showed that the coefficient for soil variable was positive and not statistically significant for the winter wheat model but negative and not statistically significant for the spring wheat model.

Similarly, the results for equation 4.17, a yield model with dummy variables for correlation date and linear time trend, showed that the coefficient for soil variable was positive and statistically significant for winter wheat and negative plus statistically significant for spring wheat. However, results for equation 4.18, a yield model with dummy variables for correlation date and linear time and quadratic time trend, showed that the coefficient for soil variable was positive and statistically significant for winter wheat but negative and not statistically significant for spring wheat. The results for equation 4.19, a yield model with dummy variables for correlation date, lag own price, linear and quadratic time trend, showed the coefficient for soil variable was positive and not statistically significant for the winter wheat model but negative and not statistically significant for the spring wheat model.

Equations 4-14 through 4.19 were also employed for total crop yield including both irrigated and non-irrigated counties. However, the results were still not significant and promising. Because of contradictory results for the wheat model, an estimation of soil survey benefits using the wheat was not conducted.

Table 5.6: Winter Wheat Yield Model Results (Publication Date)

Parameter	Estimate	Standard Error	p-value
dd_pub	0.260	0.194	0.181
trend1	0.285	0.006	<.0001
R-square	0.561	Root MSE	7.719
Coeff Var	26.024		

Parameter	Estimate	Standard Error	p-value
dd_pub	0.292	0.196	0.135
trend1	0.308	0.018	<.0001
trend2	0.000	0.000	0.174
R-square	0.561	Root MSE	7.719
Coeff Var	26.024		

Parameter	Estimate	Standard Error	p-value
dd_pub	0.249	0.199	0.210
Lag_WWprice	-1.442	0.097	<.0001
trend1	0.219	0.019	<.0001
trend2	0.000	0.000	0.241
R-square	0.565	Root MSE	7.689
Coeff Var	25.818		

Table 5.7: Winter Wheat Yield Model Results (Correlation Date)

Parameter	Estimate	Standard Error	p-value
dd_cor	0.382	0.198	0.054
trend1	0.280	0.006	<.0001
R-square	0.561	Root MSE	7.698
Coeff Var	26.130		

Parameter	Estimate	Standard Error	p-value
dd_cor	0.381	0.198	0.055
trend1	0.287	0.018	<.0001
trend2	0.000	0.000	0.666
R-square	0.561	Root MSE	7.698
Coeff Var	26.130		

Parameter	Estimate	Standard Error	p-value
dd_cor	0.186	0.202	0.357
Lag_WWprice	-1.461	0.095	<.0001
trend1	0.203	0.019	<.0001
trend2	0.000	0.000	0.047
R-square	0.565	Root MSE	7.666
Coeff Var	25.916		

Table 5.8: Spring Wheat Yield Model Results (Publication Date)

Parameter	Estimate	Standard Error	p-value
dd_pub	-0.598	0.266	0.024
trend1	0.316	0.007	<.0001
R-square	0.511	Root MSE	7.268
Coeff Var	29.643		

Parameter	Estimate	Standard Error	p-value
dd_pub	-0.449	0.275	0.102
trend1	0.368	0.025	<.0001
trend2	-0.001	0.000	0.032
R-square	0.512	Root MSE	7.266
Coeff Var	29.637		

Parameter	Estimate	Standard Error	p-value
dd_pub	-0.405	0.279	0.148
Lag_SWprice	0.001	0.127	0.994
trend1	0.352	0.028	<.0001
trend2	0.000	0.000	0.120
R-square	0.507	Root MSE	7.300
Coeff Var	29.566		

Table 5.9: Spring Wheat Yield Model Results (Correlation Date)

Parameter	Estimate	Standard Error	p-value
dd_cor	-0.933	0.215	<.0001
trend1	0.408	0.005	<.0001
R-square	0.586	Root MSE	7.098
Coeff Var	27.881		

Parameter	Estimate	Standard Error	p-value
dd_cor	-0.425	0.222	0.056
trend1	0.527	0.014	<.0001
trend2	-0.002	0.000	<.0001
R-square	0.588	Root MSE	7.077
Coeff Var	27.801		

Parameter	Estimate	Standard Error	p-value
dd_cor	-0.375	0.226	0.097
Lag_SWprice	0.413	0.096	<.0001
trend1	0.530	0.015	<.0001
trend2	-0.002	0.000	<.0001
R-square	0.582	Root MSE	7.106
Coeff Var	27.682		

5.1.4. Cotton Yield Model

Results for the upland cotton yield model using equation 4.14, a yield model with dummy variables for publication date and a linear time trend, showed that the coefficient for soil survey variable was positive and statistically significant. The results for the cotton model using equation 4.15, a yield model with dummy variables for publication date and linear and quadratic time trend, also showed that the coefficient for soil survey variable was positive and statistically significant. Likewise, results using equation 4.16 also showed that the coefficient for soil survey variable was positive and statistically significant. Similarly, results for all the equations for the correlation scenario showed that the coefficient for soil survey variable was positive and statistically significant. Table 5.10 and Table 5.11 illustrate the estimates for upland cotton model.

Equations 4-14 through 4.19 were also employed for total cotton yield including both irrigated and non-irrigated counties. However, the results were not significant and promising as for the non-irrigated portion of the cotton producing states.

Table 5.10: Upland Cotton Yield Model Results (Publication Date)

Parameter	Estimate	Standard Error	p-value
dd_pub	16.857	6.630	0.011
trend1	5.487	0.214	<.0001
R-square	0.626	Root MSE	134.397
Coeff Var	29.759		

Parameter	Estimate	Standard Error	p-value
dd_pub	31.518	7.008	<.0001
trend1	-8.417	2.226	0.000
trend2	0.125	0.020	<.0001
R-square	0.629	Root MSE	133.943
Coeff Var	29.659		

Parameter	Estimate	Standard Error	p-value
dd_pub	30.201	6.976	<.0001
Lag_cotprice	-181.812	23.933	<.0001
trend1	-2.401	2.352	0.307
trend2	0.061	0.021	0.005
R-square	0.632	Root MSE	133.277
Coeff Var	29.511		

Table 5.11: Upland Cotton Yield Model Results (Correlation Date)

Parameter	Estimate	Standard Error	p-value
dd_corr	36.658	7.494	<.0001
trend1	5.486	0.200	<.0001
R-square	0.626	Root MSE	134.436
Coeff Var	29.732		

Parameter	Estimate	Standard Error	p-value
dd_corr	25.886	7.315	0.000
trend1	-7.629	2.119	0.000
trend2	0.119	0.019	<.0001
R-square	0.623	Root MSE	134.871
Coeff Var	29.828		

Parameter	Estimate	Standard Error	p-value
dd_cor	37.503	7.455	<.0001
Lag_cotprice	-188.841	23.678	<.0001
trend1	-1.608	2.239	0.473
trend2	0.054	0.021	0.009
R-square	0.630	Root MSE	133.718
Coeff Var	29.573		

Table 5.12 shows a summary of the increase in crop yield attributable to the provision of soil information for both correlation date and publication date assumptions. The estimates for corn, soybeans, and cotton were fairly consistent when adding or removing other variables from the model. However, the estimates for both of the wheat models were not stable when adding or removing other variables from the model.

Table 5.12: Summary of Increase in Crop Yield Attributable to the Provision of Soil Information

Crop	Increase in Yield (Bushels/Acre/Year)	Model Performance
<i>Correlation Date Scenario</i>		
Corn	2.016	High R-Square and consistent estimates when adding or removing other variables, and significant
Soybeans	0.212	High R-Square and consistent estimates when adding or removing other variables, and significant
Winter Wheat	0.186	Consistent estimates but insignificant
Spring Wheat	-0.405	Negative estimates and insignificant
Upland Cotton	37.503	Consistent estimates and significant
<i>Publication Date Scenario</i>		
Corn	1.182	High R-Square and consistent estimates when adding or removing other variables, and significant
Soybeans	0.338	High R-Square and consistent estimates when adding or removing other variables, and significant
Winter Wheat	0.249	Consistent estimates but insignificant
Spring Wheat	-0.375	Negative estimates and insignificant
Upland Cotton	30.201	Consistent estimates and significant

The results for both of wheat models were not as promising as for other crop models. The reason might be the variation in management practices and spatial locations. Management practices and spatial locations for non-irrigated corn, soybeans, and cottons are homogenous compared to non-irrigated wheat. Continuous cropping and following summer fallow are the two major types of practices applied for spring wheat production.

Because of the instability and inconsistency of the estimates for the wheat models, benefit estimation of soil survey information was only considered from increased yield in corn, soybeans and cotton.

5.1.5. Validity of Fixed Effects Model

The group-mean centered method was used to determine whether fixed effects estimation results would be significantly different from the random effects results (Allison, 2005). This method suggests that if the random effects model is appropriate, which means if the time-varying independent variables are uncorrelated with county-specific fixed effects, the coefficients for the centered variables should be same as the coefficients for the mean variables. To test the hypothesis whether the random effect is uncorrelated with independent variables, the mean and deviation for each variable was calculated and then a random effects model estimated. The coefficients for means and deviations were tested for all variables together and for each variable separately. The result of group-mean centered method test is presented in Table 5.13. The coefficients and standard errors for the centered scores were similar to results using the fixed effects estimation presented in Table 5.3.

Table 5.13: Estimates from Group-Mean Centered Method

Covariance Parameter Estimates					
St. Cov Parm	Z Subject	Estimate	Error	Value	Pr > Z
Intercept	FIPS	79.335	3.956	20.060	<.0001
Residual	215.450	1.241	173.610	<.0001	

Solution for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > t
Intercept	541.350	97.924	871	5.530	<.0001
ddd_corr	2.016	0.226	60000	8.920	<.0001
dLag_CornPrice	0.584	0.134	60000	4.370	<.0001
dyear1993	-30.795	0.538	60000	-57.240	<.0001
dtrend1	0.826	0.014	60000	60.630	<.0001
dtrend2	0.009	0.000	60000	48.290	<.0001
djune_ppt	-0.368	0.141	60000	-2.610	0.009
djuly_ppt	5.547	0.161	60000	34.520	<.0001
daug_ppt	2.032	0.158	60000	12.890	<.0001
djune_mxt	-0.020	0.042	60000	-0.470	0.640
djuly_mxt	-1.417	0.048	60000	-29.230	<.0001
daug_mxt	-0.897	0.046	60000	-19.380	<.0001
djune_mnt	0.190	0.047	60000	4.030	<.0001
djuly_mnt	1.123	0.056	60000	19.900	<.0001
daug_mnt	-0.437	0.050	60000	-8.720	<.0001
mdd_corr	6.510	2.161	60000	3.010	0.003
mLag_CornPrice	-147.960	25.448	60000	-5.810	<.0001
myear1993	-31.624	120.510	60000	-0.260	0.793
mtrend1	-6.033	1.171	60000	-5.150	<.0001
mtrend2	0.057	0.015	60000	3.840	0.000
mjune_ppt	27.066	6.016	60000	4.500	<.0001
mjuly_ppt	64.020	6.895	60000	9.280	<.0001
maug_ppt	-86.068	7.837	60000	-10.980	<.0001
mjune_mxt	2.706	1.236	60000	2.190	0.029
mjuly_mxt	27.541	2.389	60000	11.530	<.0001
maug_mxt	-29.870	1.552	60000	-19.240	<.0001
mjune_mnt	8.940	1.753	60000	5.100	<.0001
mjuly_mnt	-27.873	3.708	60000	-7.520	<.0001
maug_mnt	19.785	2.272	60000	8.710	<.0001

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
all	14	60000	101.730	<.0001
dd_corr	1	60000	4.280	0.039
Lag_CornPrice	1	60000	34.070	<.0001
year1993	1	60000	0.000	0.995
trend1	1	60000	34.280	<.0001
trend2	1	60000	10.540	0.001
june_ppt	1	60000	20.790	<.0001
july_ppt	1	60000	71.880	<.0001
aug_ppt	1	60000	126.330	<.0001
june_mxt	1	60000	4.860	0.028
july_mxt	1	60000	146.820	<.0001
aug_mxt	1	60000	348.040	<.0001
june_mnt	1	60000	24.910	<.0001
july_mnt	1	60000	61.140	<.0001
aug_mnt	1	60000	79.190	<.0001

The test results (given by CONTRAST statements) show that the null hypothesis, that the deviation coefficients are same as the mean coefficients, should be rejected. Equivalently, the hypothesis that the random effect is uncorrelated with the independent variables was rejected. Tests for each individual variable showed highly significant differences for all variables except the dummy variable for the 1993 flood. The difference for the 1993 flood variable was not statistically significant. Thus the model is re-estimated by using the non-centered variables for 1993 flood variable. Re-estimating the model will show that if they were true, more efficient estimates for the coefficients will be obtained. However, there were no differences in coefficients and standard error between the output of the re-estimated model and the original group-mean centered model.

5.2. RESULTS FROM BENEFIT-COST AND SENSITIVITY ANALYSIS

The net present value (NPV) and benefit-cost ratios were calculated as part of the benefit cost analyses of the NCSS program. Benefits were estimated at the county level while costs were estimated at the state level. Benefits were evaluated only for the crop producing counties within the states included in the regressions, thus the benefits considered in this analysis are a lower bound for the total state benefits. The NPV and benefit-cost ratios were estimated for two different scenarios, the correlation date and publication date scenarios for major corn, soybean, and major cotton producing states⁵. Table 5.14 presents the sum of benefits and costs for corn and soybeans and for both correlation and publication scenarios. Table 5.15 presents the sum of benefits and costs for cotton for both correlation and publication scenarios. As presented in Table 5.14 and Table 5.15, the estimated sum of benefits was greater than the estimated total budget allocation of the soil survey program for both study regions in both scenarios.

Table 5.14: Sum of Benefits and Cost for Corn and Soybeans, 2007 Base Year

Various Discount Rate	Cost(Million Dollars)	Benefit(Million Dollars)
<i>Correlation Date Scenario</i>		
2%	\$2,030	\$24,053
4%	\$3,906	\$37,640
7%	\$11,724	\$79,456
<i>Publication Date Scenario</i>		
2%	\$2,030	\$17,647
4%	\$3,906	\$27,563
7%	\$11,724	\$54,115

⁵ See Figure 4.7, 4.8 and 4.11 in Chapter 4

Table 5.15: Sum of Benefits and Cost for Cotton, 2007 Base Year

Various Discount Rate	Cost(Million Dollars)	Benefit(Million Dollars)
<i>Correlation Date Scenario</i>		
2%	\$718	\$20,551
4%	\$1,382	\$34,198
7%	\$4,149	\$74,556
<i>Publication Date Scenario</i>		
2%	\$718	\$14,944
4%	\$1,382	\$24,612
7%	\$4,149	\$52,912

Table 5.16: NPV and Benefit-Cost for Corn and Soybeans

Various Discount Rate	NPV(Million Dollars)	Benefit /Cost
<i>Correlation Date Scenario</i>		
2%	\$22,023	11.85
4%	\$33,733	9.63
7%	\$67,732	6.78
<i>Publication Date Scenario</i>		
2%	\$15,616	8.69
4%	\$23,656	7.05
7%	\$42,390	4.62

Table 5.17: NPV and Benefit-Cost for Cotton

Various Discount Rate	NPV(Million Dollars)	Benefit /Cost
<i>Correlation Date Scenario</i>		
2%	\$19,832	28.60
4%	\$32,815	24.73
7%	\$70,406	17.97
<i>Publication Date Scenario</i>		
2%	\$14,226	20.80
4%	\$23,229	17.80
7%	\$48,762	12.75

NPV and benefit-cost ratio were calculated with respect to three different discount rates, 2%, 4%, and 7%. The results from the benefit cost analysis in terms of NPV and benefit-cost ratio are presented in Table 5.16 and Table 5.17. The benefit-cost analysis, using 7% discount rate provided an estimated benefit-cost ratio of 7:1 for the correlation date scenario and 5:1 for the publication date scenario for corn and soybeans study region. This suggests that even the lower bound estimate of benefits based on productivity increases for just two crops outweighs the cost of the entire soil survey program for the corn and soybeans study region. Similarly, benefits based on cotton productivity increases outweigh the cost of entire soil survey program for the cotton producing states included in this study.

Since it takes about three to six years to complete a soil survey, benefits were estimated for each year from 1954-2007. However, cost and NPV were estimated for each year from 1950-2007. The benefit cost analyses suggest that the net benefit in the beginning years were small and even negative for 1950-1954. Benefits were higher during 1972-1985 for corn and soybeans case as depicted in the Figure 5.1 and Figure 5.2. This is

because of the increased production during the 1970s, influenced by government farm policies and other factors, and the sharp drop in U.S. agricultural exports in early 1980s. The crops with rapidly growing exports during these periods included corn, wheat, and soybeans. The value of U.S. agricultural exports increased almost six-fold from 1970 to 1980 and the exports reached the highest peak in 1981 at \$43.78 billion⁶ (Hanrahan, 1984).

NPV and benefit-cost ratio were also estimated for each individual state. State level benefit-cost ratios for aggregate corn and soybeans production are illustrated in Table 5.18 and Table 5.19. Figure 5.3 represents the map of state-level benefit-cost ratio using a 2% discount rate for the correlation date scenario for aggregate corn and soybeans production. Iowa has the highest benefit-cost ratio of 26:1. Nebraska, Indiana, and Illinois also show high benefit-cost ratios. Missouri, Michigan and Minnesota show lower benefit-cost ratios. Iowa is the number one producer of corn and soybeans in the U.S.⁷.

⁶ This value is not adjusted for inflation.

⁷ See Table 4.6 and Table 4.7 in Chapter 4.

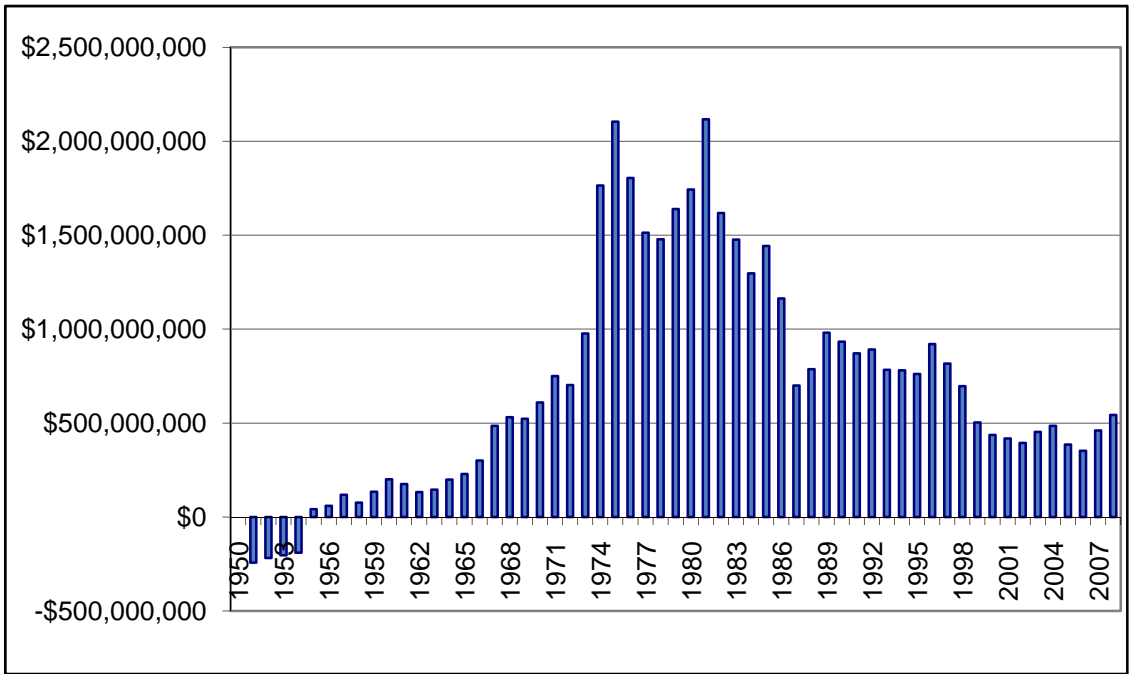


Figure 5.1: Net Benefit Graph for Corn and Soybeans

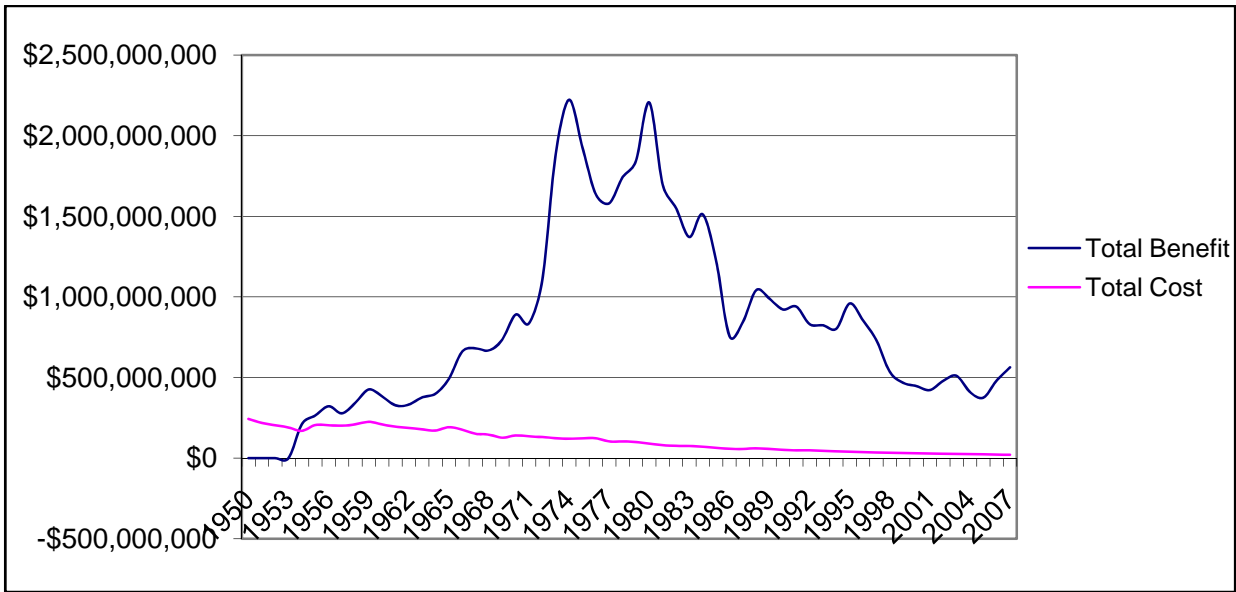


Figure 5.2: Total Benefit and Cost Graph for Corn and Soybeans

**Table 5.18: State Level Benefit-Cost Ratio for the Corn and Soybeans Region
(Correlation Date Scenario)**

State	Various Discount Rate		
	2%	4%	7%
Illinois	15.60	12.56	8.86
Indiana	17.41	13.91	9.37
Iowa	26.41	21.53	15.06
Michigan	4.58	3.81	2.78
Minnesota	7.94	6.38	4.45
Missouri	3.30	2.56	1.70
Nebraska	19.40	15.46	10.39
Ohio	13.43	11.40	8.46
South Dakota	9.09	7.58	5.73
Wisconsin	9.32	7.88	5.79

**Table 5.19: State Level Benefit-Cost Ratio for the Corn and Soybeans Region
(Publication Date Scenario)**

State	Various Discount Rate		
	2%	4%	7%
Illinois	10.99	8.72	5.69
Indiana	13.01	10.52	6.65
Iowa	19.01	15.33	9.90
Michigan	3.43	2.86	1.92
Minnesota	6.02	4.89	3.21
Missouri	2.51	1.93	1.14
Nebraska	14.11	11.14	6.95
Ohio	10.17	8.79	6.04
South Dakota	6.93	5.90	4.29
Wisconsin	6.73	5.59	3.87

State level benefit-cost ratio for cotton production is illustrated in Table 5.20 and Table 5.21. Mississippi has the highest benefit-cost ratio, with 25:1 for the correlation date scenario and 19:1 for the publication date scenario using a 7% discount rate. Mississippi and Texas produce most of the non-irrigated cotton in the U.S (Figure 4.11).

Table 5.20: State Level Benefit-Cost Ratio for Cotton Region (Correlation Date Scenario)

State	Various discount rate		
	2%	4%	7%
Arkansas	19.59	17.33	12.97
Louisiana	8.86	7.02	4.53
Mississippi	36.81	32.85	24.84
Texas	36.74	31.56	22.70

Table 5.21: State Level Benefit-Cost Ratio for Cotton Region (Publication Date Scenario)

State	Various discount rate		
	2%	4%	7%
Arkansas	11.70	10.06	7.27
Louisiana	5.49	4.19	2.56
Mississippi	28.70	25.54	19.24
Texas	27.29	23.21	16.47

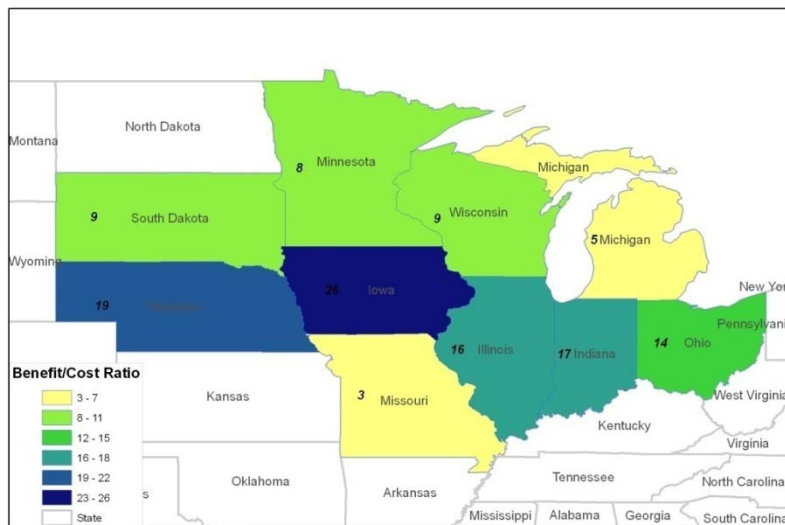


Figure 5.3: State Level Benefit-Cost Ratio for Corn and Soybeans Region (Correlation Date Scenario)

CHAPTER 6

SUMMARY AND CONCLUSIONS

6.1. SUMMARY AND IMPLICATIONS OF THE RESEARCH

This study focuses on estimating the economic benefits of the National Cooperative Soil Survey (NCSS) program in the U.S. The primary purpose of the NCSS program is to provide basic information on the soil resources of privately owned land in the US. The information provided by the soil survey program has played a significant role in such diverse fields as farming, ranching, planning, construction, and others. Soil information has long been used to guide farmers management decisions and better understand crop growth.

The NCSS program initially concentrated on the capabilities of land for agricultural production. Soil surveys provide inventories of soil resources which help farmers and other individuals to predict the potential and limitations of soils. The soil survey program provides information to farmers for site selection, land use, and management activities. Thus introduction of soil information has helped farmers to better understand and manage their land, and to make land purchase decisions.

However, because of the public good characteristics of soil survey information, the economic value of the soil survey information is not directly observed in market transactions. While a variety of non-market valuation techniques could be applied to estimate the economic value of soil information, this study relies on indirect methods to analyze econometrically estimated production relationships to infer the value of soil survey information for specific crops. The primary development is through a case study of

corn production in the Corn Belt region and an extension of the methods developed for corn to other major crops: soybeans, wheat, and cotton. Information on the implementation of the NCSS program over the past 60 years in conjunction with the data developed on crop production by USDA-NASS can be interpreted as the results of a natural experiment that can be analyzed to evaluate the effects of the soil information provided by the NCSS program on crop yield. The analysis of the information on corn production and the implementation of the NCSS program in the Corn Belt states provide robust statistical evidence of the value of the NCSS program. Applying the same approach to soybean and cotton production provides similarly robust results. The initial investigation of wheat was not as successful or robust.

In order to estimate the partial benefit of the NCSS program, crop yield models for corn, soybean, wheat, and cotton were developed and estimated using county level data from major crop producing counties. Non-irrigated counties with at least 20 years observation were included in the dataset for analysis. Corn and soybeans crop models estimated the relationship between the endogenous variable (crop yield) and exogenous variables of soil survey information, lag own price, productivity index, weather variables, and time trend. Wheat and cotton models estimated relationships between the endogenous variable (crop yield) and exogenous variables of soil survey information, lag own price, and time trend.

The econometric model used fixed effects panel data method to estimate the equations for each crop. The econometric approach relied on measuring a shift in the overall productivity for a given crop at the county level conditional on the availability of the soil information provided by the NCSS program. The best estimate of the availability

of this information is unclear. In operation, the information is released over time and some preliminary information is made available to local users while the soil survey for a specific area is being conducted. There are two specific times reported in each soil survey that can serve as proxies for the availability of information: the correlation date (the date reported in the published soil survey when the primary aspects of the information collected had been agreed upon by the soil scientists conducting the survey) and the publication date (the date of publication and full release of the NCSS report for a given area, usually a county). Thus the equations were estimated for two different scenarios based on the choice of the best time to estimate the overall effect of the soil information on crop production: (1) correlation date and (2) publication date. Since the yield of irrigated crops is influenced by the provision of irrigation, the annual production is not dependent on the same weather conditions as non-irrigated crops. Thus only data for non-irrigated crops are considered in the analysis. Since the data provided by NASS is not available for all counties for all years for all crops, only counties having at least 20 years of observations are included in the analysis.

The estimated coefficients for the provision of soil survey information were positive and highly significant for the corn, soybean, and cotton models. The results for the correlation date scenario indicate a yield increase of 2.02 bushels per acre per year for corn, 0.21 bushels per acre per year for soybeans, and 37.5 pounds per acre per year for cotton can be attributed to the provision of soil information. Similarly, the results for the publication date scenario indicate a yield increase of 1.18 bushels per acre per year for corn, 0.34 bushels per acre per year for soybeans, and 30.2 pounds per acre per year for cotton can be attributed to the provision of soil information. These results provide

substantial evidence that the soil information provided by the NCSS program has contributed to significantly increased yields for corn, soybeans, and cotton. The methods applied provide aggregate analyses that do not distinguish between the effects on the intensive and extensive margins or identify particular decisions that are influenced by the availability of additional information on the soil resources for a given area. However, it is generally agreed that soil information is used by farmers to make decisions on appropriate management practices and is a major factor in land purchase decisions.

The wheat model results were not as robust nor were the results consistent and uniformly statistically significant.⁸ This may indicate that soil information is more important for some crops (in this case non-irrigated corn, soybeans, and cotton) and not for others (in this case wheat). One reason might be variation in management practices and spatial locations.⁹ Management practices and spatial locations for corn, soybeans and cotton are more homogenous than those used for wheat. Discussions with NRCS personnel and WVU farm management extension specialist (Tom McConnell) also supported the postulation that the greater variation in management practices and spatial locations of wheat production might be the reason that the soil information coefficients for the wheat yield models were not as stable and did not indicate the same statistical stability for the importance of soil information for wheat production. This is the only known study that looks at the provision of soil information on aggregate yields. No other publications were identified that support the findings of this research that soil information is important

⁸The results from both the wheat and cotton model utilizing all the counties including non-irrigated and irrigated, were also not indicative of soil information impacts.

⁹Wheat is mainly classified as winter and spring, and among winter and spring, it is further classified as soft or hard, and white, or red (Smith, 1995). Each particular type of wheat is grown under slightly different spatial locations and climatic conditions.

for some crops such as corn, soybeans, and cotton, but not for wheat. This leaves the matter open for additional research.

In order to evaluate the economic viability of the NCSS program, two benefit-cost analyses were employed for the corn study region (Figure 4.7), the soybean study region (Figure 4.8), and the cotton study region (Figure 4.11). The first analysis is for the scenario where benefits are assumed to be accrued after the soil survey is correlated and the second analysis is for the scenario where benefits are assumed to be accrued after the soil survey information is published. An ex-post partial net-benefit of soil survey information in agricultural production is considered in this research. Benefits were estimated at the county level while costs were estimated at the state level. Benefits were considered only from the crop producing counties within the states, thus benefits considered in this analysis may not represent the total state benefits. Benefits were based on productivity gains for corn, soybeans, and cotton related to the provision of soil information at the county level. The value of increased crop yields less estimated increased production costs provides a lower-bound estimate of the economic benefits of the NCSS program. Costs were estimated based on available state level soil survey historical budget allocation and the annual budgets for the NCSS program in the U.S.

Net present values (NPV) and benefit-cost ratio were utilized to evaluate the effectiveness of the NCSS program. Aggregate county level inter-temporal benefits, in 2007 dollars of increased corn and soybean yields in the Corn Belt, were \$24 billion for the correlation date scenario and \$17 billion for the publication date scenario using a discount rate of 2%, and \$79 billion for the correlation date scenario and \$54 billion for the publication date scenario using a discount rate of 7%. Likewise, the aggregate county

level inter-temporal benefits of increased cotton yield was \$20 billion for the correlation date scenario and \$15 billion for the publication date scenario using a discount rate of 2%, and \$74 billion for the correlation date scenario and \$53 billion for the publication date scenario using a discount rate of 7%.

The benefit-cost analysis for the Corn Belt, the primary corn and soybean producing region, using a 2% discount rate estimated a benefit-cost ratio of 12:1 for the correlation date scenario and 9:1 for the publication date scenario, and using a 7% discount rate estimated a benefit-cost ratio of 7:1 for the correlation date scenario and 5:1 for the publication date scenario. The result suggests that even the lower bound estimate of benefits based on productivity increases for just two crops, corn and soybeans, outweighs the cost of the entire soil survey program for the study region. Similarly, the benefit-cost analysis for the cotton region, using a 2% discount rate estimated a benefit-cost ratio of 28:1 for the correlation date scenario and 21:1 for the publication date scenario, and using 7% discount rate estimated a benefit-cost ratio of 18:1 for the correlation date scenario and 13:1 for the publication date scenario.

This research seeks to compute a lower-bound estimate of the economic benefits of the NCSS for four major crops and thus contribute to the documentation of the value of the NCSS program soil information. The benefit-cost analyses imply that the NCSS program has provided a significant return on society's investment over the past 60 years. The results suggest that the returns of the soil survey program estimated only from increases in crop production exceed the past investment in the soil survey program in the areas considered. Given that these results indicate a substantial net benefit from only a partial use of the soil information provided by the NCSS program in major crop

production areas, the measurement of additional benefits would only increase the measured returns on societies investment. This is a promising result given the incomplete nature of the currently available data. This result combined with estimates of the value of soil information for other uses and in other sectors provides information for policy makers to make decisions on the future of the NCSS program. This study provides strong evidence that the NCSS program is viable at least in some areas of the country.

6.2. POLICY IMPLICATIONS

The results suggest that soil information has contributed significantly to increasing corn, soybean, and cotton yields in the study regions. However, the effect of provision of soil survey for wheat production was not found to be uniformly significant. The results from benefit-cost analysis suggest that the NCSS program is economically viable particularly in the corn, soybean, and cotton producing regions of the U.S. Based on empirical findings in this study even an estimate of a fraction of program benefits exceeds the total NCSS program cost. This suggests that the cost of producing the soil survey information is much lower than its benefits. The methods used in the analyses presented in this study rely on the relatively uniform production practices and data availability on output over space and time for common field crops. It is unlikely that such an approach can be generally applied. However, the analysis does imply that there are fundamental underlying factors that indicate the soil survey information has significant impacts on agricultural production. Based on the analyses presented, it seems easy to conclude investments in the NCSS program will provide significant returns over time.

The results of this research should provide information to the NRCS leadership useful in the evaluation of the NCSS program. The positive implications for aggregate

corn and soybean production suggest that further investment in developing finer scale soil maps in the Corn Belt region could be beneficial for improving yield through developments in precision agriculture. Investment in improving the county level information such as providing site specific estimates of specific soil factors, perhaps on a grid basis, could provide good overall returns to further investment.

The results derived from this research not only provide support for the NCSS program in the U.S., the implications for global development of additional soils information are evident. For example, the global project designed to provide accurate, up-to-date, and spatially referenced soil information (<http://globalsoilmap.net/>) can be expected to improve productivity on a global scale. An international consortium of soil scientists has formed with the goal of developing a new digital world soil map using currently available technologies. This research indicates that investments in such projects may provide significant benefits and can be interpreted to support and justify funding.

6.3. LIMITATIONS AND FUTURE RESEARCH

6.3.1. *Limitations*

This study provided a partial estimate of the benefits attributable to the provision of soil survey information. It employed a panel data regression approach to estimate the effects of the provision of soil survey information on aggregate production of selected primary crops and used the estimated effects to develop a benefit-cost analysis of the NCSS program. However, there are significant limitations to this study that should be improved upon in future work.

The first limitation is related to data issues for the statistical analysis. A number of theoretically relevant variables were not incorporated in the model. For example, fertilizer

and hybrid data were not included in any of the models. Despite the fact that these data would be expected to improve the performance of the model and analysis, the difficulty in obtaining such data over the time and spatial extent of the analysis limited their inclusion in this study. Omission of relevant variables could bias the results. However, time trend variables, as a proxy for technology variables, were included in the crop models to avoid the biased results that may arise because of omitted variables.

Since weather data and productivity index data were only available for corn and soybeans, these factors were not included in the wheat and cotton models. Yearly yield data for major crops are not available from a uniform starting date. For example, county level corn yield data for Kansas are only available starting in 1958, county level spring wheat yield data for Washington are only available starting in 1972, and county level cotton yield data for non-irrigated practices were only available starting in 1972. Using county yield data after the provision of the soil survey information in the analysis tends to underestimate the value of soil survey information.

The second limitation is related to benefit estimates which can be improved by further studies. Partial benefit-cost analysis underestimates the total benefits from soil surveys, particularly in states where the crops analyzed were not grown throughout the entire state. For example not all the counties in Texas, Mississippi or Louisiana produce cotton. The benefit estimates in this study are aggregated only from limited agricultural use. Only three crops (corn, soybeans, and cotton) are considered in the benefit-cost analysis. To estimate the net benefits of the NCSS program, the benefits from the soil survey to other agriculture uses and other sectors must be included. Aggregating all the benefits temporally and among different user groups is necessary to provide accurate

estimates of the net benefits of the soil survey information provided by the NCSS program.

The third limitation is related to cost estimates. The cost estimates used in this are based on the total federal budget allocation to the soil survey program. The costs numbers did not include state and/or county budget allocations, and private shares, if any¹⁰. Also, the state level federal budget estimates used in this study are approximated using only state level federal budget for 1987-1989 and the total federal budget allocation for the soil survey program. Better yearly state level budget allocations would improve the results.

The fourth limitation is the lack of information on local preferences that are important in the design and implementation of soil surveys. Local preferences for and contribution to soil surveys could have influenced the sequence of soil survey completions, such as high productivity counties may have soil surveys conducted earlier and low productivity counties later¹¹. In these cases, earlier adopters receive benefits of increased production attributed to provision of soil surveys sooner than later adopters. Exclusion of such factors in the analysis may bias the statistical estimates.

A fifth limitation relates to the aggregation of costs and benefits. Due to data limitations, costs were estimated at the state level but benefits were estimated at the county level and then aggregated to the state level. Since only counties with at least 20 years of observations were included in the analysis, for some states, the benefits are biased downward relative to costs.

¹⁰ Most of the soil survey cost is covered by federal fund, usually 80% to 90%. Only some of the states had contributed for soil survey program in past, and the contribution made is generally smaller compared to the total cost. County shares are even smaller, with some of the counties with no contribution at all.

¹¹ However, the scatter plots (Figure 4.4 to 4.6) showed that there was no significant relation between the sequences of producing soil surveys and county size, crop acreage and productivity.

6.3.2. Future Research

The limitations of this study provide an opportunity to further expand and improve the research related to valuation of the NCSS program. Future research focusing on estimating the benefits to other sectors that benefit from soil survey information is desirable.

Since the provision of the NCSS program soil information is a public good, it is difficult to capture all of the benefits that are expected to accrue. Further benefits could be estimated using other nonmarket valuation approaches. Economic tools such as survey based approaches could be useful in capturing some of the present and future benefits. Further research in this area could focus on economic benefits in other uses of soil survey information.

Some of the benefits estimation could be applied to the following sectors:

- a. Planning and construction
- b. Farming and forestry
- c. Appraisal and taxation
- d. Management and conservation

Future research study on estimating net benefits should be based on better estimates of the true program cost. Primarily, efforts should focus on obtaining better, more accurate cost information of the NCSS program.

REFERENCES

- Adams, R.M., K.J. Bryant, B.A. McCarl, D.M. Legler, J. O'Brien, A. Solow, and R. Weiher. 1995. "Value of Improved Long-Range Weather Information." *Contemporary Economic Policy* XIII:10-19.
- Adams, R. M., P. J. Farris, and D. J. Menkhous. 1983. "Response Functions and the Value of Soil Test Information: The Case of Sugar Beets." *North Central Journal of Agricultural Economics* 5(2):77-82.
- Alexander, R.B., and R.A. Smith. 1990. "County Level Estimates of Nitrogen and Phosphorus Fertilizer Use in the U.S. 1945 to 1985." U.S. Geological Survey Open-File Report 90-130. Reston, VA.
- Allison, P.D. 2005. *Fixed Effects Regression Methods for Longitudinal Data Using SAS*. SAS Publishing.
- Babcock, B.A. 1990. "The Value of Weather Information in Market Equilibrium." *American Journal of Agricultural Economics* 72(1):63-72.
- Babcock, B. A., A. L. Carriquiry, and H. S. Stern. 1996. "Evaluation of Soil Test Information in Agricultural Decision-Making." *Applied Statistics* 45:447-61.
- Baltagi, B. 1995. *Econometric Analysis of Panel Data*. New York: Wiley.
- Bean, L.H. 1967. "Crops, Weather, and the Agricultural Revolution." *The American Statistician* 21(3):10-14.
- Beckett, P. H. T., and P.A. Burrough. 1971. "The Relation between Cost and Utility in Soil Survey: IV. Comparison of the Utilities of Soil Maps Produced by Different Survey Procedures and to Different Scales." *Journal of Soil Science* 22(4):466-480.
- Beckett, P. H. T. 1981. "The Cost-Benefit Relationships of Soil Surveys." In *Soil Resource Inventories and Development Planning*, Tech. Mono., 1, Soil Management Support Service, USDA.
- Berlage, L., and R. Renard. 1985. "The Discount Rate in Cost-Benefit Analysis and the Choice of a Numeraire." *Oxford Economic Papers*, New Series 37(4):691-699.
- Bie, S.W., and P.H.T. Beckett. 1971. "Quality Control in Soil Survey. II. The Costs of Soil Survey." *Journal of Soil Science* 22(4):453-465.
- Bie, S. W., A. Ulph, and P.H. T. Beckett. 1973. "Calculating the Economic Benefits of Soil Survey." *Journal of Soil Science* 24(4):429-435.
- Bie, S. W., and A. Ulph. 1972. "The Economic Value of Soil Survey Information." *Journal of Agricultural Economics* 23:285-297.

- Braunstien, Y. 1981. "Information as a Commodity: Public Policy Issues and Current Research." In R. M. Mason and J. E. Creps, eds. *Information Services: Economics, Management and Technology*. Boulder CO: Westview, pp. 9-22.
- Brown, M. J. 1959. "The Relation of Weather Factors to the Yield of Winter Wheat in Box Elder County, Utah." *Monthly Weather Review* 87:97-99.
- Buller, O. 1972. "Influence of Research and Policy on Crop Yields in Kansas." *Transactions of the Kansas Academy of Science* 75(1):20-28.
- Chavas, J., and R. D. Pope. 1984. "Information: Its Measurement and Valuation." *American Journal of Agricultural Economics* 66(5):705-710.
- Costello, C.J., R.M. Adams, and S. Polasky. 1998. "The Value of El Niño Forecasts in the Management of Salmon: A Stochastic Dynamic Assessment." *American Journal of Agricultural Economics* 80:765-777.
- Dent, D., and A. Young. 1981. *Scale, Accuracy, Costs and Returns*. London: George Allen & Unwi.
- Ditzler, C.A., R.J. Engel, and R.J. Ahrens. 2003. "Soil Taxonomy and Soil Survey." In H. Eswaran, R. J. Ahrens, T. J. Thomas, and B. Stewart, eds. *Soil Classification: A Global Desk Reference*. Washington, DC: CRC Press, pp.221-230.
- Dobos, R. R., H. R. Sinclair, and K. W. Hipple, cited 2008. User Guide National Commodity Crop Productivity Index (NCCPI), Version 1.0 [Available on line at ftp://ftp-fc.sc.egov.usda.gov/NSSC/NCCPI/NCCPI_user_guide.pdf] accessed on May 26, 2008
- Durana, P.J., and D. Helms.2002. "Soil Survey Interpretations: Past, Present, and Looking to the Future." In P. J. Durana, A. B. Effland and D. Helms, eds. *Profiles in the History of the U.S. Soil Survey*. Ames: Iowa State Press, pp.275-302.
- Freeman, A. M. III. 1993. *The Measurement of Environmental and Resource Values*. Washington, DC: Resources for the Future.
- Garcia, P., S.E. Offutt, M. Pinar, and S.A. Changnon. 1987 "Corn Yield Behavior: Effects of Technological Advance and Weather Conditions." *Journal of Applied Meteorology* 26:1092-1102.
- Giasson, E., C. van ES., and A. van Wambeke. 2000. "Assessing the Economic Value of Soil Information Using Decision Analysis Techniques." *Soil Science* 165(12):971-978.
- GlobalSoilMap.net. Cited 2009 [Available online at <http://globalsoilmap.net>] accessed on April 22, 2009
- Greene, W. 2003. *Econometric Analysis*, 5th.ed. Prentice Hall: Englewood Cliffs.

- Griliches, Z. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica* 25:501-522.
- Gujarati, D. 2004. *Basic Econometrics*, 4th. ed. McGraw-Hill.
- Hanrahan, C. E. 1984. Why U.S. Agricultural Exports Have Declined in the 1980s. Congressional Research Service, Library of Congress. Washington D.C., USA [Available online at <http://digital.library.unt.edu/govdocs/crs/permalink/meta-crs-8863>] accessed on April 10, 2009
- Heimlich, R. E. 1986. "Agricultural Programs and Cropland Conversion, 1975-1981." *Land Economics* 62(2):174-181.
- Hilton, R. W. 1981. "The Determinants of Information Value: Synthesizing Some General Results." *Management Science* 27(1):57-64.
- Hirshleifer, J. 1973. "Where Are We in the Theory of Information?" *American Economic Review* 63:31-39.
- Houck, J. P., and P. W. Gallagher. 1976. "The Price Responsiveness of U.S. Corn Yields." *American Journal of Agricultural Economics* 58:731-34.
- Hsiao, C. 2003. *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- Hu, Q., and G. Buyanovsky. 2003. "Climate Effect on Corn Yield in Missouri." *Journal of Applied Meteorology* 42:1626-1635.
- Huff, F. A., and J. C. Neill. 1982. "Effects of Natural Climatic Fluctuations on the Temporal and Spatial Variation in Crop Yields." *Journal of Applied Meteorology* 21:540-550.
- Just, R., D. L. Hueth, and A. Schmitz. 2004. *The Welfare Economics of Public Policy: A Practical Approach to Project and Policy Evaluation*. Cheltenham, UK: Edward Elgar.
- Kaufmann, R.K., and S. Snell. 1997. "A Biophysical Model of Corn Yield: Integrating Physical and Economic Determinants." *American Journal of Agricultural Economics* 79(1):178-180.
- Kaylen, M. S., J. W. Wade, and D. B. Frank. 1992. "Stochastic Trend, Weather, and U.S. Corn Yield Variability." *Applied Econometrics* 24:513-18.
- Klingebiel, A.A. 1966. "Costs and Returns of Soil Survey." *Soil Conservation* 32:3-6.
- Lang, H. J., and D. N. Marino. 1993. *The Selection Process for Capital Projects*. New York: John Wiley & Sons, Inc.
- Lawrence, D. 1999. *The Economic Value of Information*. New York: Springer-Verlag.

- Lave, L. 1963. "The Value of Better Weather Information to the Raisin Industry." *Econometrica* 31(2):151-163.
- Leblanc, D. C. 2004. *Statistics: Concepts and Applications for Science*. Jones & Bartlett Publishers.
- Lubowski, R. N., S. Bucholtz, R. Claassen, M. J. Roberts, J. C. Cooper, A. Gueorguieva, and R. Johansson. 2006. "Environmental Effects of Agricultural Land-Use Change: The Role of Economics and Policy" Economic Research Report No. (ERR-25) [Available online at <http://www.ers.usda.gov/publications/err25/err25.pdf>]
- Macauley, M. K. 2005. "The Value of Information: A Background Paper on Measuring the Contribution of Space-Derived Earth Science Data to National Resource Management." Discussion Paper 05-26. Washington, DC: Resources for the Future.
- Manogaran, M.C. 1981. "Assessing the Relative Effects of Climate and Technology on Wheat Production in Kansas." *Transactions of the Kansas Academy of Science* 84 (2):65-77.
- McCall, J. J., ed. 1982. *The Economics of Information and Uncertainty*. Chicago: University of Chicago Press.
- McGee, J., and L. Prusak. 1993. *Managing Information Strategically*. New York: John Wiley & Sons, Inc.
- Menz, K. M., and P. Pardey. 1983. "Technology and U.S. Corn Yields: Plateaus and Price Responsiveness." *American Journal of Agricultural Economics* 65(3):558-562.
- Meyer, B. D. 1995. "Natural and Quasi-experiments in Economics." *Journal of Business & Economic Statistics*, American Statistical Association 13(2):151-61.
- Michigan State Chief Climatologist's Office. cited 2008. [Available online at <http://climate.geo.msu.edu/>]
- Mitchell, R. C. and R. T. Carson. 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.
- Offutt, S. E., P. Garcia, and M. Pinar. 1987. "Technological Advance, Weather, and Crop Yield Behavior". *North Central Journal of Agricultural Economics* 9(1):49-63.
- Paul, D. M. 2003. "Value of Imperfect Input Information in Agricultural Production." *Agricultural Systems* 75(2-3):277-294.
(<http://www.sciencedirect.com/science/article/B6T3W-47MHWDG-9/2/4865fee69175b039f209b5d0f4fcf983>)
- Perrin, R. K. 1976. "The Value of Information and the Value of Theoretical Models in Crop Response Research." *American Journal of Agricultural Economics* 58(1):54-61.

- Preckel, P., E. Loehman, and M. Kaylen. 1987. "The Value of Public Information for Microeconomic Production Decisions." *Western Journal of Agricultural Economics* 12(2):193-197.
- Radner, R., and J. Stiglitz. 1984. "A Nonconcavity in the Value of Information," In: M. Boyer, and, R. Kihlstrom, eds. *Bayesian Models in Economic Theory*. New York: Elsevier Science Publishers, pp. 33-52.
- Reddy, V.R., and D.N. Baker. 1990. "Application of GOSSYM to analysis of the effects of weather on cotton yields." *Agriculture Systems* 32:83-95.
- Repo, A. J. 1989. "The Value of Information - Approaches in Economics, Accounting, and Management Science." *Journal of the American Society for Information Science* 40(2):68-85.
- Roe, T., and F. Antonovitz. 1985. "A Producer's Willingness to Pay for Information Under Price Uncertainty: Theory and Application." *Southern Economic Journal* 52:382-91.
- Rosenzweig, C., A. Iglesias, X.B. Yang, P. R. Epstein and E. Chivian. 2001. "Climate Change and Extreme Weather Events; Implications for Food Production, Plant Diseases, and Pests." *Global Change & Human Health*. 2(2): 90-104 (<http://chge.med.harvard.edu/about/faculty/journals/rosenzweig2001.pdf>)
- Samuelson, J. R., M. Stelford, and R. J. Daniel. 2002. "The Importance and Value of Soil Information." 2002 ASAE Annual Meeting, Paper number 021093. St. Joseph, Michigan www.asabe.org
- Schlenker, W., and M.J. Roberts. 2006. "Nonlinear Effects of Weather on Corn Yields." *Review of Agricultural Economics* 28:391-398.
- Schroder, D., J.C. Headley, and R. M. Finley. 1984. "The Contribution of Herbicides and Other Technologies to Corn Production in the Corn Belt Region, 1964 to 1979." *North Central Journal of Agricultural Economics* 6(1):95-104.
- Shen, S.S.P., P. Dzikowski, G. Li, and D. Griffith. 2001. "Interpolation of 1961-97 Daily Temperature and Precipitation Data onto Alberta Polygons of Ecodistrict and Soil Landscapes of Canada." *Journal of Applied Meteorology* 40:2162-2177.
- Simonson, R.W. 1955. "Changing Place of Soils in Agricultural Production." *The Scientific Monthly* 81(4):173-182.
- Smith, D. W. 1998. "Soil Survey During Its Infancy of the Late 1890's and Early 1900's" National Cooperative Soil Survey Centennial Highlights. Papers for the 1998 Annual Conference of the Soil and Water Conservation Society, Presented July 1998, San Diego, California [Available online at <ftp://ftp-fc.sc.gov.usda.gov/NSSC/NCSS/History/1998swcpapers.pdf>]
- Smith, C.W. 1995. *Crop Production*, New York: John Wiley & Sons, Inc.

- Soil Survey Division Staff. 1993. Soil Survey Manual. Soil Conservation Service. U.S. Department of Agriculture Handbook 18. Available online at <http://soils.usda.gov/technical/manual/> accessed on January 19, 2009.
- Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at <http://websoilsurvey.nrcs.usda.gov/> accessed on April 22, 2009
- Soil Survey Staff, Soil Conservation Service, National Soil Survey Handbook, title 430-VI (Washington D.C. U.S. Government Printing Office, November 1993)
- Solow, A. R. Adams, K. Bryant, D. Legler, J. O'Brian, B. McCarl, W. Nayda, and R. Weiher. 1998. "The Value of ENSO Forecasts: The Case of U.S. Agriculture." *Climate Change* 39:47-60.
- Stiglitz, J. E. 2000. "The Contributions of the Economics of Information to Twentieth Century Economics." *Quarterly Journal of Economics* 115(4):1441-1478.
- Swanson, E.R., and J.C. Nyankori. 1979. "Influence of Weather and Technology on Corn and Soybean Yield Trends." *Agricultural Meteorology* 20:327-342.
- Swinton, S. M. and R. P. King. 1994. "The Value of Pest Information in a Dynamic Setting: The Case of Weed Control." *American Journal of Agricultural Economics* 76(1):36-46.
- Thompson, L.M. 1969. "Weather and Technology in the Production of Corn in the U.S. Corn Belt." *Agronomy Journal* 61:453-456.
- Thompson, L.M. 1970. "Weather and Technology in the Production of Soybeans in the Central U.S." *Agronomy Journal* 62:232-236.
- Thompson, L.M. 1985. "Weather Variability, Climatic Change, and Soybean Production." *Journal of Soil and Water Conservation* 40:386-389.
- Thompson, L.M. 1988. "Effects of Changes in Climate and Weather Variability on the Yields of Corn and Soybeans." *Journal of Production Agriculture* 1:20-27.
- U.S. Bureau of the Budget 1950- 2009. Budget of the U.S. Government.
- USDA-ERS Cited 2008. [Available online at <http://www.ers.usda.gov/data/majorlanduses/spreadsheets/croplandusedforcrops.xls>] accessed on April 12, 2008
- USDA-NRCS Cited 2008 [Available online at <http://soildatamart.nrcs.usda.gov/StatusMaps/SoilDataAvailabilityMap.pdf>] accessed on April 24, 2009
- USDA- NASS, 1997. Crop Production 1996 Summary. [Available on line at <http://usda.mannlib.cornell.edu/usda/nass/CropProdSu//1990s/1997/CropProdSu-01-00-1997.pdf>] assessed on September 24, 2008

- USDA- NASS, 2008. Crop Production 2007 Summary. [Available on line at <http://usda.mannlib.cornell.edu/usda/current/CropProdSu/CropProdSu-01-11-2008.pdf>] assessed on September 24, 2008
- USDA-NASS Cited 2008. Quick Stats: Agricultural Statistics Data Base. [Available online at http://www.nass.usda.gov/QuickStats/Create_County_All.jsp]
- USDA-NASS Cited 2008. Quick Stats: Agricultural Statistics Data Base. [Available online at http://www.nass.usda.gov/QuickStats/PullData_US.jsp]
- USDA-NRCS Cited 2009 [Available online at <http://www.nrcs.usda.gov/programs/>] accessed on Jan 29, 2009
- USDA-NRCS Cited 2009 [Available online at http://gcmd.nasa.gov/records/USDA_NRCS_WSS.html] accessed on March 22, 2009
- USDA-NRCS. 2000. Soil Survey of Ozark County, Missouri
- USDA-Soil Conservation Service. 1988. Soil Survey of Brown County, Illinois
- USDC-BEA Cited 2009. [Available online at <http://www.bea.gov/national/#gdp>] accessed on Feb 08, 2009
- U.S.Office of Management and Budget. 1992. *Circular A-94: Discount Rates to be Used in Evaluating Time-Distributed Costs and Benefits.*
- Western, S. 1978. "Survey Quality and Survey Value." In: S. Western, *Soil Survey Contracts and Quality Control*. 284f. Monographs (Soil Survey) - Oxford University.
- Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT.
- Wooldridge, J. M. 2006. *Introductory Econometrics: A Modern Approach*, 3rd. ed. Thomson South-Western.
- Young, A. 1973. "Soil Survey Procedures in Land Development Planning." *Geographical Journal* 139:53-64.
- Young, R. A. 2005. *Determining the Economic Value of Water: Concepts and Methods*. Washington, DC: Resources for the Future.
- Zink, N.E. 1940. "The Relation of Weather Factors to Wheat Yields on Levan Ridge, Utah." *Monthly Weather Review* 68:66-71.

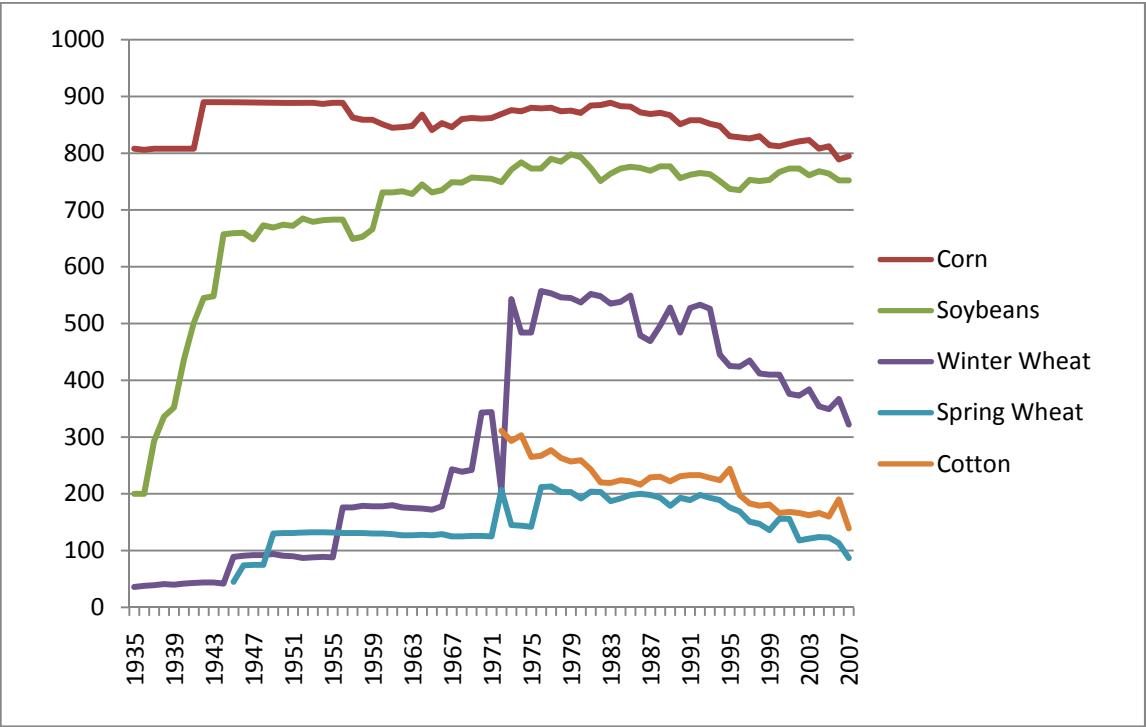
APPENDIX A

Table: Number of Crop Data Available Each Year for the Study Region

YEAR	Corn	Soybeans	Winter Wheat	Spring Wheat	Cotton
1935	808	200	36	77	
1936	806	200	38	77	
1937	808	293	39	77	
1938	808	336	41	77	
1939	808	352	40	77	
1940	808	436	42	77	
1941	808	501	43	76	
1942	890	545	44	77	
1943	890	548	44	77	
1944	890	657	42	76	
1945	890	659	89	121	
1946	890	660	91	151	
1947	889	648	92	151	
1948	889	673	92	151	
1949	889	669	94	204	
1950	889	674	91	204	
1951	889	672	90	204	
1952	889	685	87	204	
1953	889	679	88	204	
1954	887	682	89	204	
1955	889	683	88	204	
1956	889	683	176	204	
1957	863	649	176	182	
1958	859	653	179	188	
1959	859	666	178	202	
1960	851	731	178	201	
1961	845	731	180	202	
1962	846	733	176	195	
1963	848	728	175	198	
1964	868	745	174	196	
1965	841	731	172	191	
1966	853	735	178	193	
1967	846	749	243	198	
1968	860	748	239	199	
1969	862	757	242	199	
1970	861	756	343	199	

YEAR	Corn	Soybeans	Winter Wheat	Spring Wheat	Cotton
1971	862	755	344	199	
1972	869	749	204	265	311
1973	876	771	543	207	293
1974	874	784	484	210	303
1975	880	773	484	209	265
1976	879	773	557	269	267
1977	880	790	553	269	277
1978	874	785	546	265	263
1979	875	798	545	266	257
1980	871	793	537	256	259
1981	884	774	552	266	243
1982	885	751	548	259	220
1983	889	764	535	243	219
1984	883	773	538	244	224
1985	882	776	549	265	222
1986	872	774	479	265	216
1987	869	769	469	246	229
1988	871	777	496	244	230
1989	867	777	528	224	222
1990	851	756	484	229	231
1991	858	762	527	223	233
1992	858	765	533	230	233
1993	852	763	526	234	228
1994	848	751	445	230	224
1995	830	737	425	215	244
1996	828	735	424	205	198
1997	826	753	435	195	183
1998	830	751	412	184	179
1999	814	753	410	173	181
2000	812	767	410	192	166
2001	817	773	376	192	168
2002	821	773	373	155	166
2003	823	761	384	158	162
2004	808	768	354	160	166
2005	812	764	349	161	160
2006	789	752	367	154	190
2007	795	752	322	135	139

Graph showing Number of Crop Data Available Each Year for the Study Region



APPENDIX B

Table: List of Excluded States from Top 10 Crop Producing States in 2007

Crop Model	Excluded States	Reasons for not included in the analysis
Corn	Kansas	Kansas and Michigan usually are at similar ranking for corn production, however USDA-NASS yield data are available earlier for Michigan (starting in 1942) than for Kansas(starting in 1958)
Soybean	North Dakota Arkansas	Crop production rank for Missouri is higher than for North Dakota and Arkansas based on both 2007 and 1996 crop production summary. USDA-NASS data are available earlier for Michigan (starting in 1942) than for Arkansas (starting in 1947 for some counties, and later for others), and some of the counties in Arkansas are irrigated.
Winter Wheat	Idaho	Most of the counties in Idaho are irrigated.
Spring Wheat	Oregon Colorado Utah Wisconsin	Most of the counties in these states are irrigated.
Cotton	Georgia North Carolina Missouri California Tennessee Arizona	Most of the counties in these states are irrigated.

APPENDIX C

Corn Yield Reduced Model (Correlation Date Scenario)

R-Square	Coeff Var	Root MSE	YIELD Mean
0.840655	20.77805	16.30976	78.49511

Parameter	Estimate	Standard Error	t Value	Pr > t
Cor_dum	2.16988904	0.25072944	8.65	<.0001
Lag_Cornprice	-1.25375166	0.14233174	-8.81	<.0001
year1993	-22.03691715	0.57130251	-38.57	<.0001
Trend	1.17140168	0.01370509	85.47	<.0001
Trend2	0.00525156	0.00018069	29.06	<.0001

R-Square	Coeff Var	Root MSE	YIELD Mean
0.839078	21.15023	16.45331	77.79260

Parameter	Estimate	Standard Error	t Value	Pr > t
Cor_dum	1.710768833	0.25177370	6.79	<.0001
Trend	1.165271103	0.01289743	90.35	<.0001
Trend2	0.005569923	0.00017690	31.49	<.0001

Corn Yield Reduced Model (Publication Date Scenario)

R-Square	Coeff Var	Root MSE	YIELD Mean
0.840775	20.73247	16.31074	78.67247

Parameter	Estimate	Standard Error	t Value	Pr > t
Pub_dum	2.33916286	0.25411850	9.21	<.0001
Lag_Cornprice	-1.22122551	0.14257563	-8.57	<.0001
year1993	-21.91819654	0.57081904	-38.40	<.0001
Trend	1.19061993	0.01373866	86.66	<.0001
Trend2	0.00502628	0.00018756	26.80	<.0001

R-Square	Coeff Var	Root MSE	YIELD Mean
0.839230	21.10162	16.45308	77.97069

Parameter	Estimate	Standard Error	t Value	Pr > t
Pub_dum	2.038901878	0.25549085	7.98	<.0001
Trend1	1.180567418	0.01293770	91.25	<.0001
Trend2	0.005342820	0.00018374	29.08	<.0001

Soybeans Yield Reduced Model (Correlation Date Scenario)

R-Square	Coeff Var	Root MSE	YIELD Mean
0.797806	16.79068	4.797508	28.5724

Parameter	Estimate	Standard Error	t Value	Pr > t
Cor_dum	0.390485512	0.08082966	4.83	<.0001
Lag_Soyprice	-0.243382246	0.02316651	-10.51	<.0001
year1993	-2.248030042	0.17767659	-12.65	<.0001
Trend1	0.382345735	0.00537420	71.14	<.0001
Trend2	0.000488017	0.00006839	7.14	<.0001

R-Square	Coeff Var	Root MSE	YIELD Mean
0.798104	16.84610	4.802352	28.5072

Parameter	Estimate	Standard Error	t Value	Pr > t
Cor_dum	0.3019401556	0.08064943	3.74	0.0002
Trend1	0.3611958544	0.00479312	75.36	<.0001
Trend2	0.0007761272	0.00006069	12.79	<.0001

Soybeans Yield Reduced Model (Publication Date Scenario)

R-Square	Coeff Var	Root MSE	YIELD Mean
0.800778	16.64316	4.761437	28.6089

Parameter	Estimate	Standard Error	t Value	Pr > t
Pub_dum	0.576479462	0.08089012	7.13	<.0001
Lag_Cornprice	-1.350232863	0.04941666	-27.32	<.0001
year1993	-2.502614189	0.17635050	-14.19	<.0001
Trend1	0.331966674	0.00495772	66.96	<.0001
Trend2	0.000704136	0.00006325	11.13	<.0001

R-Square	Coeff Var	Root MSE	YIELD Mean
0.798513	16.80578	4.796926	28.54332

Parameter	Estimate	Standard Error	t Value	Pr > t
Pub_dum	0.5421373141	0.08134860	6.66	<.0001
Trend	0.3642810069	0.00475666	76.58	<.0001
Trend2	0.0006872990	0.00006280	10.94	<.0001

APPENDIX D

Corn Yield Model Results with NCCPI (Publication Date)

R-Square	Coeff Var	Root MSE	YIELD Mean
0.871055	18.65861	14.67919	78.6724

Parameter	Estimate	Standard Error	t Value	Pr > t
dd_pub	1.81168019	0.22895911	7.91	<.0001
NCCPI	0.00000000 B	.	.	.
Lag_Cornprice	0.60625448	0.13377822	4.53	<.0001
trend1	0.83931139	0.01364975	61.49	<.0001
trend2	0.00871380	0.00018731	46.52	<.0001
june_ppt	-0.36897124	0.14111071	-2.61	0.0089
july_ppt	5.54718069	0.16089223	34.48	<.0001
aug_ppt	2.02422606	0.15786176	12.82	<.0001
june_mxt	-0.01217108	0.04222231	-0.29	0.7731
july_mxt	-1.41325057	0.04856746	-29.10	<.0001
aug_mxt	-0.89450450	0.04638504	-19.28	<.0001
june_mnt	0.18378367	0.04711354	3.90	<.0001
july_mnt	1.11483119	0.05648993	19.74	<.0001
aug_mnt	-0.44834306	0.05022876	-8.93	<.0001

(NOTE: The X'X matrix was been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.)