Economic Aspects of Crop Pest Management and Monarch Conservation

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Tiesta Thakur

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Terrance M Hurley, Advisor

August, 2020

Acknowledgement

The life of a graduate student is definitely not a cakewalk. The perennial existential crises and self-doubts are your constant companions. In those moments of crisis, to have a team of experienced researchers who not only believe in you but also religiously mentor you is a bountiful blessing. I have been extremely fortunate to have not one but many helping hands to keep me going through all my highs and lows. I am very grateful to my advisor, Dr. Terrance Hurley for his expert guidance and dedication to my academic and professional development. His acuteness and more importantly calm disposition helped me maintain my perspective and focus on the bigger picture, which is very easily lost once you are deep into your research work.

I am also thankful to Dr. Frances Homans, Dr. Robert Haight, Dr. Joseph Ritter and Dr. Alan Benson with whom I had the privilege to collaborate on different research projects. Every one of those experience helped me understand the making of an authentic and successful researcher. Special thanks to Dr. Robert Koch for his indispensable advice on the entomological aspect of my research work.

Most of all I want to thank my husband, Kaushalendra Kishore, for our countless brainstorming sessions which never let me miss the minute yet important details.

To those who made it possible but will never read it.
Mom, Dad, Ishan, Vihaan and KK

Abstract

The dramatic rise in prophylactic chemical treatments for pest and weed control in the past two decades have raised many environmental concerns. Although cost effective, such treatments which include neonicotinoids and herbicide tolerant (HT) crops contaminate soil and water and cause wildlife habitat loss. In my thesis, I explore the economics of eco-friendly practices of the farmers ranging from bio-diversity conservation to adapting integrated pest management. In the first chapter, I survey Midwestern farmers to estimate their willingness to grow milkweed on their non-cropland for Monarch butterfly conservation for various remuneration rates. I also approximate intrinsic motivation of farmers from their actual conservation data using reverse regression, distance discriminant analysis and control functions to test for motivation crowding out. Findings indicate motivation crowding out at modest levels of compensation. Alternatively, high remuneration crowds in farmers motivation to conserve Monarchs. The second chapter estimate soybean farmers' value of information provided by alternative configurations of a monitoring network for soybean rust (*Phakopsora pachyrhizi*). It shows that a network of 400 sentinel plots can maximize the expected profit of soybean farmers provided more plots are placed in the Corn Belt where the risk of soybean rust infection is lower, but where much more soybean is produced in contrast to the current spatial arrangement where sentinel plots are dis-proportionately placed in the Southern US. The last chapter examines the economic suitability of Unmanned Aerial Vehicles (UAVs) for scouting soybean aphids (*Aphis glycines* Matsumura) based on a plant-level spatiotemporal bioeconomic model of infestation. Findings indicate that the optimal profit from UAV based scouting is equivalent to that from manual scouting. But its greater tendency to detect false positives can also trigger frequent unnecessary treatments and dramatically reduce farmers' profits. Yet, UAV's commercial viability depends more on reducing its operating cost than improving its precision, once it has a tally threshold of 250 soybean aphids per plant.

Contents

List of Figures v

List of Tables vi

1 Introduction 1

2 Will Farmers Grow Weeds for monarchs? 3

2.1 Introduction . . . 3
2.2 Survey Design . . . 6
2.3 General Setup of the model . . . 8
2.4 Results . . . 18
2.5 Conclusion . . . 26

3 Valuing Monitoring Networks for New Pathogens: The Case of Soybean Rust 29

3.1 Introduction . . . 29
3.2 Farmer Decision Model . . . 31
3.3 Dynamic Optimization . . . 37
3.4 Farm Model Implications . . . 39
3.5 Spatial Optimization Model . . . 41
3.6 Results . . . 42
3.7 Conclusion . . . 48

4 Will Farmers Adopt Remote Sensing for Soybean Aphid Management? An Economic Perspective. 50

4.1 Introduction . . . 50
4.2 Literature Review . . . 53
4.3 Bioeconomic Model . . . 56
4.4 Model Parametrization . . . 64
4.5 Simulating the model . . . 67
4.6 Results . . . 68
4.7 Conclusion . . . 72

Bibliography 74

Appendices 88

A.1 . . . 89
A.2 . . . 95
B.1 . . . 105
B.2 . . . 107
C.1 . . . 108
C.2 . . . 112

List of Figures

2.1 Average conditional probability of voluntary participation for farmers who volunteered or got paid (r_j) given their total number of conservation activities (N_j). . . . 11
2.2 IM1.2- Probability of voluntary participation weighted by w_j^1 12
2.3 (a) Distances d_j^H, d_j^L and IM2.1, and (b) IM2.2 as a function of farmers total number of conservation activities and their decision to get paid for them. . . . 15
2.4 IM Proxies, $\hat{U}_j^I$, based on $\hat{U}_j = -D1$ and its average given the total number of conservation activities N_j and payment status r_j. . . . 17
2.5 Percentage of respondents willing to participate in monarch conservation program as a function of payment offered. . . . 20
2.6 The effect of compensation level on the average probability of participation in monarch conservation for quartiles of IM-1.2. . . . 24
2.7 The effect of compensation level on the average probability of participation in monarch conservation for every quartile of IM-2.2. . . . 25

3.1 Optimal management decisions for a farmer with no monitoring network. 40
3.2 (a) Correlation function ρ and (b) optimal management decisions for a farmer with monitoring network when $\phi^f = \phi^s$. . . . 41
3.3 Optimal strategies when (a) monitoring network is unavailable and (b) post-spatial optimization when B=80. . . . 43
3.4 Value of monitoring (in millions) when each county has a sentinel plot. 44
3.5 Optimal sentinel plot locations when (a) B=185 and (b) B=80. . . . 45
3.6 Optimal strategies when (a) monitoring network is unavailable and (b) post-spatial optimization when B=80. . . . 46
3.7 Counties with (a) sentinel plots in 2014 and (b) sentinel plots post-spatial optimization when B=39. . . . 46
3.8 Value of monitoring by number of sentinel plots (budget). . . . 47

4.1 Fitted sigmoidal proportional yield loss function versus linear loss function. 66
4.2 Expected profit in \$ per hectare using (a) Manual scouting and (b) Remote sensing given ϕ equals 250 aphids per plant. 68
4.3 Percentage yield loss and total number of treatments as a function of economic thresholds (a) ψ for manual scouting and (b) κ for remote sensing given $\phi = 250$. 69
4.4 (a) Isoprofit lines and (b) Optimal ET for UAV based scouting as a function of probability of type 1 and 2 errors. 71
4.5 Relationship between the ratios of expected optimal profits $E(\pi_u^*)/E(\pi_m^*)$ and monitoring costs C_u/C_m when C_m is fixed at \$9.88 ha^{-1}. 71
A.1.6 Transformation plots for CATPCA in Method II. 89
A.1.7 Scree plots for CATPCA in Method II. 90
A.1.8 Component scores color coded based on total number of conservation activities and type of conservation payment. The axes D1 and D4 are the principal components from CATPCA in Method II. 90
A.1.9 Scree plots for CATPCA in Method III. 91
A.1.10 Transformation plots for CATPCA in Method III. 92
A.1.11 Distribution for IM proxies. 93
A.1.12 Average number of programs farmers participate in and percentage of farmers paid by quartiles for all IM proxies. 94
B.1.1 Probability of infection. 105
B.1.2 (a) Soybean acreage and (b) yield (bu/acre) in 2014. 105
C.1.1 Population N_t as a function of t when $N_{max} \sim N(1000, 100)$, $T_{max} \sim N(36.29, 0.42)$ and $b \sim N(0.228, 0.23)$. 108

List of Tables

2.1 Summarizing regression results from IM proxies based on methods I, II and III. 22
2.2 Effect of interaction between IM proxies 1.2, 2.2 and 3.2, and getting paid on the probability of participation in the monarch conservation program. 27

3.1 States of the world and their probability of occurrence when ipmPIPE is available. 33

A.2.1 Probit regression using dummy variable for payments as the dependent variable. 95

A.2.2 Component Loadings for the first five principal components, D1 to D5, from CATPCA in Method II. 95
A.2.3 Component loadings from CATPCA done in Method III. 96
A.2.4 Linear Regression where different functions of -D1, the first principal component, are used to approximate the dependent variable. 96
A.2.5 Summary statistics 97
A.2.6 Farm management practices of farmers. 97
A.2.7 Conservation Practices of farmers. 98
A.2.8 Average difficulty level of requirements by the monarch habitat establishment types. 98
A.2.9 Probit Regression with participation in the monarch conservation program as the dependent variable. 99
A.2.10 Probit regression with participation in the monarch habitat conservation program as the dependent variable and intrinsic motivation proxies 1.1 and 1.2. 100
A.2.11 Probit regression with participation in the monarch habitat Conservation program as the dependent variable and IM proxies 2.1 and 2.2. . 101
A.2.12 Probit Regression with participation in the monarch habitat conservation program as the dependent variable and IM proxies 3.1 and 3.2. . 102
A.2.13 Effect of interaction between IM proxy and getting paid on the probability of participation in the monarch conservation program. 103
A.2.14 Attributes of the monarch habitat conservation programs. 104
B.1.1 Baseline parameters for the farmer decision model. 105
B.1.2 Summary Statistics 106
C.1.1 Transition rules for the state variables of a cell i at time t. 108
C.1.2 Percentage yield loss as a function of the probabilities of type 1 and 2 error. 109
C.1.3 Parameters and their values. 109
C.1.4 Optimal expected profit from UAV based scouting, $E(\pi_u^*)$, as a function of the probabilities of Type 1 and 2 error in \$ per hectare. 110
C.1.5 The likelihood of treatment as a function of the probabilities of type 1 and 2 error. 111

Chapter 1:
Introduction

The turn of the century witnessed a huge surge in prophylactic crop management in the United States ranging from the widespread adoption of herbicide tolerant (HT) and insect resistant (IR) crops to neonicotinoid seed treatments. They have significantly reduced the weed and pest management cost and consequently enhanced farmers profitability on average. But their excessive usage have not only led to rapid evolution of resistant pests and weeds but also aggravated the risk of bio-diversity degradation. A 2012 survey of crop growers in 31 states showed a 34% rise in the glysophate resistant weeds since 2010 [77]. Evidences on pyrethroid resistance among many major commodity crop pests such as soybean aphids [80] and western corn rootworm [140] have also been piling up. Neonicotinoids which are the most used insecticides among the U.S. crop growers in terms of area treated [67], are associated with the declining population of several pollinators. Henry *et al.* [56] show that non-lethal exposure of honey bees to neonicotinoids can damage their navigational skills which can put their colony at risk of collapse. The increase in toxicity of milkweeds by Clothianidin [117] and the adverse effect of HT crops on milkweed population [113] have all contributed to the significant drop in the monarch butterfly population in recent decades. All these incidences point out the misalignment in the farm management goals of the farmers and the society. The central theme of my thesis is to understand the economics of pro-environment farm practices of the farmers ranging from their involvement in bio-diversity conservation to adopting integrated pest management.

In my first chapter, I survey farmers in the Midwest and estimate their willingness to participate in hypothetical monarch habitat restoration programs with varying levels of financial assistance. I also approximate intrinsic motivation of farmers from their actual conservation data using reverse regression, distance discriminant analysis and control functions to test the motivation crowding out hypothesis. The results show that farmers experience motivation crowding out and consequently lower average participation at modest levels of remuneration. The crowding out is predominant among farmers whose estimated intrinsic motivation are low on average. Predictably, they also experience significant crowding in of motivation at high levels of remuneration. On the contrary, participation decisions of farmers with high estimates of intrinsic motivation, on average, vary less with the conservation payments. Overall the findings suggest that a volunteer program will not only be cost effective but also more successful.

The second chapter examines producers' value of information provided by alternative configurations of a monitoring network for an invasive wind-borne crop disease, soybean rust (*Phakopsora pachyrhizi*). Soybean producers use monitoring information to understand how exposed their farms are to soybean rust infection and to make fungicide application decisions. The value of information is the expected gain in producers' profit from using the monitoring network, and it is estimated using a dynamic model of producer decision-making. Findings indicate that the value of the sentinel plot network increases with the number of sentinel plots, reaching a maximum at 400 sentinel plots. In addition, the optimal spatial arrangement of sentinel plots generated by our model is substantially different from the actual placements. Current sentinel plots are disproportionately placed in the Southern US where the risk of infection is high, but the amount of soybean is relatively small. Our estimates suggest more plots should be placed in the Corn Belt where the risk of an infection is lower, but where much more soybean is produced.

The third chapter analyses the economic suitability of Unmanned Aerial Vehicles (UAVs) for scouting soybean aphids based on a plant-level agent based pest management model. This model generates optimal economic thresholds and treatment strategies for farmers when they scout their field manually and when they use UAVs. We find that manual scouting triggers insecticide treatment in the third and fourth weeks since the arrival of soybean aphids once the optimal threshold of 275 aphids per plant is reached. Aerial surveillance will lead to an insecticide application in the fourth week once 45% of the plants are infested with at least 250 aphids per plant. Both techniques generate equivalent optimal expected profits. But because UAVs are more prone to Type II error, they can trigger frequent unnecessary treatments. As a result, manual scouting is a more suited to farmers facing low risk of soybean aphid infestation and also to those who prefer an early prophylactic treatment. Yet, to make UAV's a commercially viable scouting technique, it is more important to reduce its operating cost than improve its precision, once it has a tally threshold of 250 soybean aphids per plant.

Chapter 2:

Will Farmers Grow Weeds for monarchs?

I estimate farmers' willingness to participate in hypothetical monarch habitat restoration programs with varying levels of financial assistance based on a survey of Midwestern farmers. I also approximate intrinsic motivation of farmers from their actual conservation data using reverse regression, distance discriminant analysis and control functions and test the motivation crowding out hypothesis. The results confirm motivation crowding out and consequently lower average participation at modest levels of remuneration, particularly for the farmers who have limited conservation history. Predictably, high compensation rates increase their average participation significantly. Conversely, the farmers who already practice several types of conservation, especially volunteers, are more likely to participate in the monarch conservation program, without caring too much for the proposed financial incentives.

Keywords: Monarch conservation, milkweed, intrinsic motivation, motivation crowding out.

2.1 Introduction

In the past two decades, monarch butterflies have experienced a significant decline in their population [134],[1] so much so, that they are being considered for listing under the Endangered Species Act of 1973. Based on the annual monitoring of their overwintering sites in Mexico, this area has declined at the rate of 0.89ha/year [138]. Researchers believe that the large scale reduction in milkweed, which is monarch larvae's critical food source, plays a major role in aggravating the issue at hand. For most of the twentieth century, a large majority of monarchs in the eastern U.S. originated from the milkweed population in agricultural fields [113] which since the introduction of herbicide tolerant (HT) corn and soybean have substantially declined. In fact, 98% of the loss in milkweed is attributed to the loss of milkweed in soybean and corn fields [120]. As a result, both the Government and private agencies in the U.S. are taking initiatives to restore monarch habitat. The U.S. Fish and Wildlife Services plans to restore sufficient habitat

Support for this project was provided by BASF Corporation through the grant "Farmer Perceptions of the Living Acres Initiative & the Cultivation of 'Beneficial Weeds'." and the Minnesota State Agricultural Experiment Station (Project MIN-14-134).

[1] Semmens *et al.* [134] report that the monarch population has declined by 84% between the winters of 1996-97 and 2014-15, and face a quasi-extinction of 11-57% in next 20 years.

by 2020 to support a monarch population occupying 6 ha of overwintering habitat in Mexico [147]. Pleasants [120] estimated that 425 million milkweeds need to be added to increase the monarch support capacity by one more overwintering hectare. This means that in order to meet the 6 ha conservation goal, a total of 1.6 billion milkweed should be planted.

Milkweed, although important for beneficial insects and greater biodiversity, are generally unwanted due to their negative impact on crop production. This presents a dilemma about how to maintain such weeds without compromising crop production. One possible solution to this problem is growing "beneficial weed" in patches on the farm landscape. With such a concept, four relevant questions are: i) Where is the best location on the farm landscape for beneficial weed patches? ii) What is the best way to establish and maintain them? iii) Who will be responsible for establishment and maintenance? and iv) Are there more efficient ways to secure the benefits they can provide? Each of these questions has important socioeconomic as well as biological dimensions. For example, farmers may be unwilling to maintain weeds within their farm fields even if it is the most productive area from a biological perspective. Alternatively, while farmers maybe ready to broadcast weed seed into ditches and fence rows to establish beneficial weed patches, germination success may be too poor to make it worthwhile from a biological perspective. Consequently, it is essential to identify motivators, extrinsic or otherwise, that will make farmers agreeable to planting the socially optimal level of beneficial weeds.

The objective of this research was to begin to determine the extent to which farmers would be willing to establish and maintain milkweed patches on the farm landscape to promote monarch butterfly habitat. This objective was accomplished using a survey to assess farmers' preferences over three hypothetical monarch habitat restoration programs on the farm landscape under a range of volunteer and financial incentives. Specifically, monarch habitat could be established either using milkweed seeds and peat pots, root plugs or by broadcasting seed mix of native milkweed, grasses and other wildflower. Two of the programs were designed around BASF Corporation's Living Acres #Monarch Challenge,[2] while the third was designed around USDA Conservation Reserve and Stewardship Program requirements. The restoration could be voluntary or for a onetime payment of $100, 250, 500 or 1000 per acre. Every questionnaire contained a combination of one establishment method and financial incentive. The results of the survey indicate the following: (i) Farmers' average willingness to participate range from 33.3% to 65.8% depending on the proposed incentives and farmer's state of operation. (ii) Those agreeing to participate would enroll 2.45 acres of their non-cropland into the program on average. The level of enrolment depends on her expectation about the program's effectiveness: the likelihood of producing enough milkweed per acre and percentage of other participating farmers [75]. (iii) Farmers' participation is not significantly different across the establishment methods given the level of monetary payments.

[2] agriculture.basf.com/us/en/Crop-Protection/Living-Acre.html

But, on average they do find broadcasting to be the most convenient method. (iv) Farmers would rather volunteer than receive modest financial remuneration [48, 78], thereby hinting at motivation crowding out. Estimates suggest that financial incentives in excess of $610 per acre would be required to get participation that is higher than from a purely volunteer program.

Payment for Environmental Services (PES) have been consistently used in many countries to promote biodiversity and ecosystem services conservation [10, 83, 20, 105]. In the United States, National Resources Conservation Service (NRCS), through the Environmental Quality Incentives Program (EQIP), provides financial and free technical assistance to the agricultural producers in Midwest and Southern Great Plains, to plant milkweed and other nectar rich plants on private lands. Besides the direct financial gains, PES can also crowd in intrinsic motivation which indirectly helps in attaining the desirable outcome. PES sometimes symbolize acknowledgement of pro-social behaviour and consequently enhance the self esteem of individuals [150, 151, 139]. It can also have a "prescriptive effect" where it signals the individuals about what constitutes socially desirable actions [12, 130]. PES is also reported to crowd in intrinsic motivation in common pool resource games [8, 137, 112] especially when it facilitates interpersonal communication which in turn foster mutual trust among the participants. But PES is also known to backfire. Based on the Self-Determination Theory, PES may crowd out people's motivation if it is perceived as a signal of distrust or inhibits their autonomy, competence, social and environmental relatedness [30, 37]. In fact, in the recent years, evidence in support of it crowding out intrinsic motivation has significantly increased [42, 48, 112, 27, 46], more so, than its anti-theory (crowding in). Rode *et al.* [129] reviewed eighteen empirical studies published since 2013 on motivation crowding by economic incentives in conservation policy and found thirteen of them in support of crowding out hypothesis. In this paper, I explore the possibility of motivation crowding out by onetime payments for monarch conservation. Testing motivation crowding hypothesis is especially important when designing a new conservation plan because PES effects can extend to post-PES intervention time period [8]. In their study on conservation activities of the rural communities in Mexico, Garcia-Amado *et al.* [46] find that longer PES intervention can erode people's inherent and cultural interest in conservation in favor of utilitarian interest.

To test the crowding out hypothesis, I first measure the intrinsic motivation of farmers to conserve based on their current conservation activities. Most popular methodologies for testing motivation crowding theory include framed field experiments, role games and ethnographic approach. These techniques are more prone to social desirability and hypothetical bias and therefore can have limited external validity [109, 9]. But here, since intrinsic motivation is evaluated using farmers' actual conservation behavior data, the results are more robust to any of the above mentioned biases. Three different data reduction techniques are employed on farmers' conservation data to approximate their intrinsic motivation. These methods are related to the concepts of reverse regression, distance based discriminant analysis and control functions. Farmer's intrinsic motivation (IM) is expected to be non-decreasing in her number of conservation activities. It should also be higher for those who conserve voluntarily rather than for payment. These two properties serve as the exclusion criteria for my IM proxies. The regression results show that payments, on average, crowd out intrinsic motivation of farmers to conserve monarchs, particularly at low levels. Yet, high levels of conservation payment successfully increase probability of participation, especially for those who have low IM scores. On the contrary, participation decisions of farmers with high estimates of IM, on average, vary less with the conservation payments. Overall a volunteer program will not only be cost effective but also more successful.

The structure of the paper is as follows. In section 2.2, I describe the survey. The basic results from the survey are summarized in subsection 2.4.1 of section 2.4. Section 2.3 first discusses the general set up of the econometric model and its theory. It explains the motivation behind conservation practices, both intrinsic and extrinsic, their role in encouraging conservation and crowding out effects. Next, the techniques for approximating intrinsic motivation are elaborated in sections 2.3.1, 2.3.2 and 2.3.3. All the proxies should uphold some basic properties of intrinsic motivation outlined in section 2.3. Finally, they are used in a regression framework to test the theory of crowding out of intrinsic motivation. The regression results are reported in subsection 2.4.2 of section 2.4. Section 2.5 concludes.

2.2 Survey Design

A farmer survey was conducted in Minnesota, Iowa and Wisconsin. Its objective was to give insight into the farm management and conservation practices of the farmers in general as well as to estimate their willingness to participate in monarch conservation. The questionnaire is organized into 4 sections covering different aspects of farmers' conservation and farm management practices. The survey had a single bounded, dichotomous choice design to keep it simple and avoided anchoring bias. Section I focuses on understanding the farm management activities of the farmers. They are asked whether they practice activities such as tillage, insecticide treatments, scouting for pests, herbicides, use of herbicide tolerant (HT) seeds, cover crops, and rotated crops to manage their farm. It also contains questions about farm size, whether it is organic certified or not, crops planted and planted acres for each crop. Section II focuses on the conservation activities of the farmers. Farmers are questioned whether they have ever participated in USDA conservation programs such as Conservation Reserve Program (CRP), Environmental Quality Incentive Program (EQIP), Agricultural Conservation Easement Program (ACEP), Conservation Reserve Enhancement Program (CREP) and Wildlife Habitat Incentives Program (WHIP). They are also asked if they are involved in conserving soil, water, pollinator, wildlife or prairie and whether they received any monetary, technical or other type of assistance for the same.

Section III describes a hypothetical 5 year monarch conservation program to the farmers for establishing milkweed habitat on their non-crop land. The program has two main attributes: (i) method of establishment and (ii) monetary payment for participation. There are 3 methods of establishment, each characterized by a unique set of requirements, which are then paired with 5 levels of monetary payments, thus making it a total of 15 unique programs. The 3 types of establishment programs considered are as follows:

1. Establishment using milkweed seeds and peat pots
2. Establishment using milkweed root plugs
3. Establishment with seed mix containing 5% native milkweed, 15% native grasses and 80% native wildflower

The inputs will be provided to the farmers free of cost. The milkweed patch, irrespective of the establishment type, should receive daily a minimum of 6 hours of sunlight and not be treated with insecticides. For method I, milkweed seeds should be sowed in the peat pots 4 weeks before the last expected frost. The seedlings should be kept moist and indoors, with daily exposure to sun up until the danger of frost has passed. Then, they should be transplanted into the desired location and planted at least 15 feet apart. In method II, the milkweed patch is established using root plugs. The root plugs should be planted at least 15 feet apart and 4-6 inches deep to avoid excessive drying from exposure to air. In both of these methods, any surrounding vegetation should be removed before transplanting milkweeds to avoid competition [36] and plants should be checked to make sure that they have adequate moisture. In case rainfall is not sufficient, milkweed should be watered.

When adopting establishment method III, cold stratification of seeds can prove beneficial. Planting in fall when the first frost starts to occur is one of the most traditional strategies for cold stratification. It is because at this time of the season, soil is mostly wet and does not need additional watering. The location where establishment is to be made must be treated with non-selective herbicides twice, once in the fall and again in spring, to clear any competing vegetation. The seeds should be mixed with carriers such as cat litter, sawdust and soy hulls in order to facilitate good seed coverage. The seed mix should then be broadcasted at a rate of 50 pure live seed per square foot. It is important to make sure that the seed bed is firm. To do so, the site can be rolled with rollers or drive across with truck.

Once milkweed is planted (using any of the above methods), mowing should be done monthly during the first year in order to curb the growth of competing vegetation. In the coming 2-5 years, no herbicides or insecticides should be used on the monarch habitat. To ensure a healthy habitat, mowing should be repeated in the fall of the third year. The programs and their attributes are summarized in Table A.2.14.

All the establishment programs offer farmers one of the following monetary payments for every acre they commit to the program.

1. $0 or Voluntary participation

2. One time establishment payment of $100 per acre

3. One time establishment payment of $250 per acre

4. One time establishment payment of $500 per acre

5. One time establishment payment of $1000 per acre

Each questionnaire includes one establishment method and level of monetary payment. It describes the program in detail to the survey participant and then asks if she is willing to participate in it. Interested farmers are further queried about how many acres and where on their non-cropland will they establish the monarch habitat. Farmers are also asked what factors, both in general and particular to the conservation program, motivate them to participate. These include how easy they find complying to the attributes of the program, their understanding of cultural and environmental benefits (e.g., pollination) of monarch butterflies, availability of technical and financial assistance and time, to name a few. Section IV of the survey is dedicated to the demographics of the respondents.

The survey was conducted in two phases using mail and mixed modes. The mixed mode technique included mailing questionnaires to those who did not respond to an initial online surveys and reminder. In the beginning, postcards were sent out to 3600 farmers explaining the purpose of the survey. Phase I began a week later when questionnaires were sent out to the farmers, half mail and half email. After a week, participants who received mail and online surveys were sent reminders through postcards and emails respectively. In the following week, Phase II was initialised and questionnaires were mailed to all those who had not responded after Phase I.

2.3 General Setup of the model

Farmer j participates in a conservation program i when her utility from participation, denoted by U_{ji}, is greater than her opportunity cost C_i. Let y_{ji} denote her participation decision, such that $y_{ji} \in \{0, 1\}\ \forall i, j$. U_{ji} is made up of the intrinsic and extrinsic utilities, labeled U^I_{ji} and U^E_{ji} respectively. Suppose U_{ji} is additive in its components [42]:

$$U_{ji} = U^I_{ji} + U^E_{ji} \quad \forall i, j. \tag{2.1}$$

While there are various types of extrinsic motivators ranging from Government regulations to peer pressure [23, 43], in this paper I focus exclusively on conservation payments.

Let r_{ji} denote farmer *j's* decision to accept money for conservation. It takes value 1 if farmer j participates in program i in exchange for money $z_i (> 0)$ and 0 otherwise. Both U^E_{ji} and U^I_{ji} are functions of $z_i \; \forall \; i$. On one hand, extrinsic or financial utility U^E_{ji} will always increase with z_i. Assume that U^E_{ji} is a linear function of z_i [42, 18]. On the other hand, U^I_{ji} can be increasing or decreasing in z_i depending on whether monetary rewards crowd in or out intrinsic motivation (IM) [130, 139, 78]. It can also be non-monotonic in z_i [48]. Let N_j denote the total number of conservation activities of farmer j ($N_j = \Sigma_i y_{ji}$). In the survey, farmers were only asked about their conservation practices related to soil, water, pollinator, wildlife and prairie. Therefore, $N_j \in \{1, 2, 3, 4, 5\} \; \forall \; j$. Their total utility from N_j programs, labeled U_j, is then an aggregate of their utilities U_{ji} derived from every program i they participate in, i.e. $U_j = \Sigma_i y_{ji} U_{ji}$. In the rest of the paper, the aggregate concepts are denoted in the same way as their counterpart, except the subscripts for the dimensions being aggregated will be suppressed. I also lose the subscript i whenever referring to monarch conservation specifically.

My objectives are to understand (i) how U^I_j and U^E_j affect farmer *j's* decision to conserve monarchs and (ii) whether U^I_j is crowded out by payment offered for conservation. To accomplish these goals, U^I_j and U^E_j must be measured. Suppose their proxies for any conservation program i are $\hat{U}^I_{ji}$ and $\hat{U}^E_{ji}$ respectively. Since U^E_{ji} is strictly monotonic in the dollar amount $z_i \; \forall i$, the dollar amount offered to establish milkweed will make a good $\hat{U}^E_j$. U^I_j is unobserved and will be estimated based on the techniques discussed below. Finally, I estimate the following equation:

$$f(y_j) = \alpha_0 + \alpha_1 \hat{U}^I_j + \alpha_2 g(z) + \beta X_j + \tau X_i + \epsilon_j. \tag{2.2}$$

The dependent variable in Equation 2.2, $f(y_j)$, and the regressor *g(z)* are monotonic functions *'f'* and *'g'* of participation dummy y_j and dollar amount z respectively. X_j and X_i are other person specific and program specific covariates . Because y_j is binary, either ordinary least square, probit or logistic regression can be used to estimate Equation 2.2.

In this paper, farmer's intrinsic motivation for conservation is assessed based on their N_j and $r_j \; \forall \; j$. Given my understanding of intrinsic motivation, I assume that it will have the following properties:

- It is non-decreasing in $N_j \; \forall \; j$, keeping the payment level fixed.
- Given N_j, conserving voluntarily implies greater intrinsic motivation $\forall \; j$.

Any good IM proxy should also share these properties. The dictum that 'past behavior is the best predictor of future behavior' is supported by much empirical evidence in the literature on psychology [4, 5]. Farmers intrinsic motivation to conserve can be approximated using their current conservation practices related to soil, water, pollina-

[3] See Table A.2.7 for details on conservation activities of the farmers in Appendix A.2.

tor, wildlife and prairie and the monetary payments they received, annual or onetime, for their participation.[3] I also only predict farmers' aggregate utility from conservation instead of their program specific utilities. In order to keep the questionnaires short, farmers were not asked particulars of their conservation practices. For example, they were not questioned about the dollar amount they received, if any, for any of their conservation activities. They were, however, asked if they got paid for any of their conservation practices. As a result, U_j^E can be approximated using $r_j = I(\Sigma_{i=1}^{N_j} r_{ji} > 0)$ where I is an indicator variable which equals 1(0) when farmer j reports (volunteering) taking money.

Note that the conservation behaviour of people are bound to be highly correlated because they are all determined by some common underlying factors such as their IM. So, in the process of estimating farmers' intrinsic motivation to conserve, I perform dimension reduction on their conservation data to minimize multicollinearity. Hereinafter I discuss the different methodologies which were used to construct IM proxies.

2.3.1 Method I

As already discussed above, the intrinsic utility from conservation should be greater for those who are doing it voluntarily. In method I, I estimate how likely farmer j is to volunteer in one of the conservation programs. The binary variable, r_j, is regressed on the participation decision variables, y_{ji} for all i and some other characteristics of farmer j. r_j equals 1 if the farmer j reports accepting monetary reward, annual or onetime, for her conservation activities. The regression equation can be stated as follows:

$$P(r_j = 1) = \delta_0 + \Sigma_i \delta_i y_{ji} + \psi X_j + e_j. \tag{2.3}$$

Here e_j is the residual term. y_{ji}'s include the dummy variables for participation in soil, water, pollinator, wildlife and prairie conservation. X_j contains log of farmer j's farm size relative to the median farm size in her agricultural district, her education level, age and agricultural district fixed effects. Table A.2.1 summarizes the regression results using the probit model.[4] It shows that practicing any type of conservation, except soil, has a higher chance of getting monetarily rewarded. Relatively large farms are also more likely to get paid for participating in conservation activities. In Figure 2.1, the turquoise lines represent the average probability of farmers volunteering for all their conservation activities, $E_j(P(r_j = 0))$, grouped by number of conservation activities and their payment level. They are negatively sloped because volunteering in multiple conservation activities is costlier than volunteering for just one. Consequently, volunteering becomes less likely when the number of conservation activities increase. Furthermore, the predicted probability is always higher for a farmer j who volunteered to conserve than farmer j' who was paid for participating, given that $N_j = N_{j'}$.

[4] Additional tables and figures for this chapter are in and named after appendix A and are duly numbered.

The participation decision variables y_{ji}'s and the dependent variable $P(r_j = 1)$ both depend on intrinsic motivation of farmer j which is not explicitly controlled for in Equation 2.3. Hence, my regression may suffer from omitted variable bias. Moreover, 2.3 is a reverse regression because participation in conservation activities may depend on monetary assistance and not vice versa. The direction of bias in coefficients δ_i will depend on whether farmers, on average, experience crowding in or out of intrinsic motivation when paid for conservation. But, the purpose of estimating 2.3 is data reduction, i.e. to express the chance of voluntary participation as a linear combination of conservation activities of the farmers, and not to establish causality. That is why, I do not focus on statistical inferences of the results.

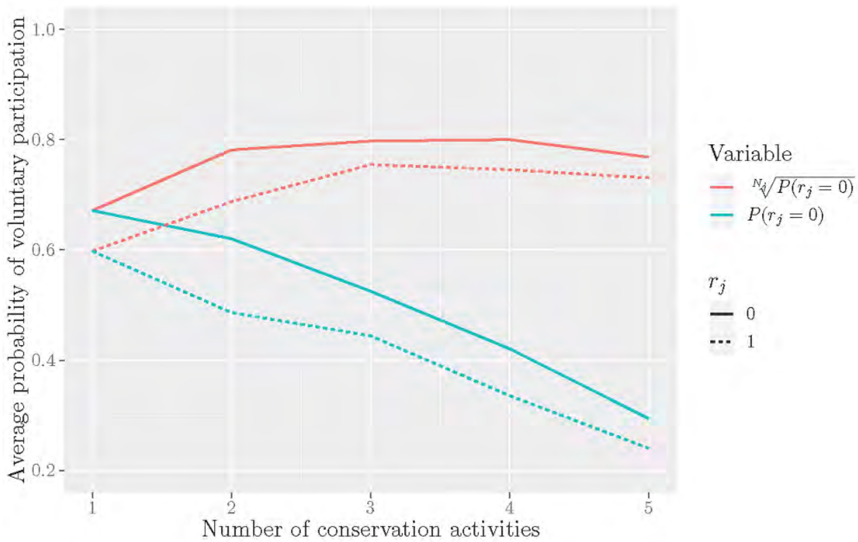


Figure 2.1: Average conditional probability of voluntary participation for farmers who volunteered or got paid (r_j) given their total number of conservation activities (N_j).

Suppose that farmers' decision to practice different kinds of conservation are independent of each other after controlling for their intrinsic and extrinsic motivation. Then, their probability of volunteering in a conservation program will be $\sqrt[N_j]{P(r_j = 0)}$. This probability is represented by the red lines in Figure 2.1. Ideally, the ex-post probability of volunteering in only one conservation program should be higher for the volunteers with greater N. But, Figure 2.1 shows that it is not non-decreasing in the total number of activities for volunteers. The average probability of volunteering rises from 0.66 to 0.8 when total conservation activities increase from one to three. It then decreases to 0.796 and 0.75 when the latter is four and five respectively. These two means, however, are not equal only at 11% level of significance.

One of the reasons for such anomaly could be the diminishing sample size of volunteers with an increase in conservation activities. Their number is reduced four times (131:62 to 10:18) in comparison to those getting paid when total activities expand from one to five. It is fair to say that these numbers imitate the real world scenario where not many farmers practice multiple types of conservation and that too pro-bono. But, it can potentially bias $\sqrt[N_j]{P(r_j = 0)}$ which systematically loses its variance with diminishing sample size. Its standard deviation declines continuously with the number of conservation practices from 0.13 when N_j=1 to 0.079 at N_j=5. The deteriorating variance could have been checked by incorporating more variables in Equation 2.3. But getting additional information would have made the questionnaire lengthy which in turn could have dropped the response rate [40, 132]. Instead, I control for the bias by weighing the proxy variable with the inverse of its standard deviation conditional on the number of programs.[5] Let weight w_j^1 equal

$$w_j^1 = \frac{1}{\sqrt{Var(\sqrt[N_j]{P(r_j = 0)})|N_j)}} \quad \forall j.$$

Multiplying w_j^1 by $\sqrt[N_j]{P(r_j = 0)}$, transforms it monotonically and consequently makes the weighted proxy relatively more non-decreasing in N_j than its unweighted counterpart. Figure 2.2 shows the average of the IM proxy, scaled by w_j^1, as a function of farmer's total conservation activities N_j and their decision to take money for their services r_j. I refer to $\sqrt[N_j]{P(r_j = 0)}$ and $w_j^1 \sqrt[N_j]{P(r_j = 0)}$ as IM-1.1 and IM-1.2 respectively.

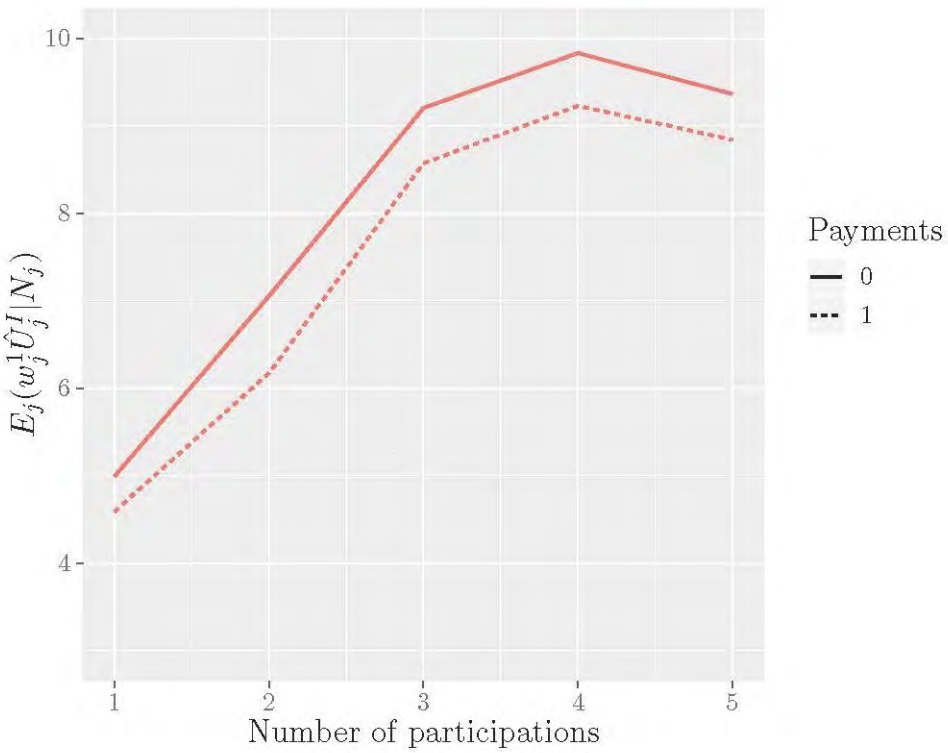


Figure 2.2: IM1.2- Probability of voluntary participation weighted by w_j^1

[5]I use variance and not probability of occurrence to construct the weights due to two reasons. First, as already stated, it is not unexpected for the sample size to decrease with increase in number of conservation activities. Nevertheless, it is hard to know whether the sample size for each N fairly represent the underlying population. Second, while the same logic holds in case of variance, yet it is more likely that variance based on small samples are underestimates of their true values. Farmers

2.3.2 Method II

Here, farmers are assigned relative IM indices with respect to the most and least intrinsically motivated to conserve. To do so, I first find those two dimensions (principal components) of the data which are dominated by N_j and r_j using Categorical Principal Component Analysis (CATPCA). Subsequently, farmers are represented by their component scores on these two principal components. Section 2.3 clearly states that intrinsic motivation is non-decreasing in N_j and decreasing in r_j. Therefore, as per the survey data, group of farmers with N_j=5 and r_j=0 have the highest environmental morale. On the other hand, those practicing only one conservation plan (N_j=1) and that too with financial assistance (r_j=1) form the least motivated group.[6] Let the centroid of these two groups be represented by 'H' and 'L' respectively. Furthermore, suppose farmer j is at a euclidean distance of d_j^H and d_j^L from H and L respectively. Then, those who are highly (less) intrinsically motivated to conserve will have low (high) d_j^H and high (low) d_j^L.

PCA is one of the most widely used data reduction techniques [108, 3]. But, it is only suited for data sets with continuous variables [90]. Categorical Principal Component Analysis (CATPCA), however, optimally quantifies qualitative variables and then performs the linear PCA on the transformed variables [90, 103, 104]. Optimal scaling or quantification is nothing but nonlinear or linear transformation of variables which maximize their correlation. It is however a constrained optimization. The admissibility of transformation functions depend on whether the variables are nominal, ordinal or numeric. For a nominal analysis level, the optimal category quantification does not need to be monotonic. In case of ordinal and numerical variables however, the optimal transformation preserves the rank-order of the original variables. The transformed numerical variables should also maintain the original relative spacing of the underlying variables. Here, CATPCA is performed on farmers' conservation practices related to soil, water, pollinators, prairie and wildlife, and their age, education level and farm size relative to the median farm size in their agricultural district. Other demographic variables such as race and gender are not incorporated in CATPCA because they produce negligible variation in the data. Information on their conservation payments are also included. As discussed before, this survey asks farmers whether they received monetary payments, annual or onetime, for their conservation activities. Therefore, every farmer belongs to one of these 4 categories: They are either voluntary participants, paid annually, onetime or both. I effectively consolidate these information into one nominal variable, named 'Payment type', which takes value 0, 2, 3 and 4 for the categories above and use it for

taste for conservation is determined by multiple factors, many of which are unobserved. So, it is expected to vary more than what the sample estimates for any N.

[6] In the survey, farmers are asked to check the boxes for what all conservation they practice. Therefore, those who did not check any of the boxes could be either be farmers who did not practice any conservation at all or those who chose to not respond. These two types can be very distinct from each other. The survey did not have questions which could distinguish between the two. That is why, I define the least motivated group of farmers as those who practice at most one conservation instead of none.

analysis. The variables are quantified during dimension reduction based on the following set of restrictions- all conservation practices and variable for types of payment received are treated nominally while education and N are ordinal. The rest of the variables are continuous and therefore are classified as numerical. I also impose rank-1 restriction on quantification matrix to keep the analysis simple [104].

Preliminary results from CATPCA are shown in figures A.1.7 and A.1.6. Figure A.1.6 shows the optimally transformed variables plotted against their observed values. The plotted transformation is with respect to the first principal component.[7] As already mentioned, for a nominal variable, the optimal category quantification does not need to be monotonic. This is very clear in the case of variable 'Payment type'. In case of ordinal and numerical variables, the optimal transformation preserves the rank-order of the original variables. For example, the age of farmers are scaled so that older farmers continue to get higher score on the optimally transformed age variable. Similarly, people who participated in more conservation activities (N) receive higher value on the transformed N as well.

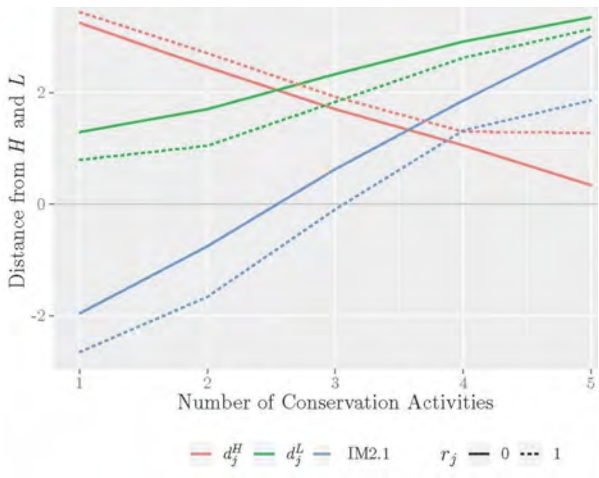
Figure A.1.7 is the scree plot with the first five principal components. These dimensions account for 27.7%, 15.52%, 10.83%, 10.01% and 9.87% of the total variance respectively. Table A.2.2 gives the principal component loadings of the variables on dimensions 1 ('D1') to 5 ('D5'). The results show that the most dominant factors for D1, D2, D3, D4 and D5 are N, log of farmers' farm size relative to the median farm size in their agricultural district, payment type, age and education respectively. Furthermore, pollinator, wildlife and prairie conservation are highly correlated with D1 and weak on D2 unlike soil conservation which greatly influence D2 without making much difference to D1. To construct the IM proxy, I will focus on D1 and D4 only because the two dimensions capture the most variability in N and information on conservation payments.

Farmer j can be represented by her coordinates $\{D1(j), D4(j)\}$ ∀ j. The small dots in Figure A.1.8 represent farmers on {D1, D4} coordinate system. They are color coded based on combinations of farmers total number of activities, N_j, and whether they got paid for it, r_j, making it a total of ten groups. The big dots are the centroid for each group. Farmers who participate in more conservation activities score low on D1. On the other hand, those who conserve voluntarily score high on D4. Therefore, the centroid of groups H and L are located in the second and fourth quadrant. Farmers in the second (fourth) quadrant who are inevitably closer to (farther from) H and farther from (closer to) L have higher (lower) IM to conserve. But then there are those, like the ones in the first and third quadrants, who are far away from both H and L. Their distance estimates d_j^H and d_j^L independently do not provide good estimates of their intrinsic motivation. That is why, I use $d_j^L - d_j^H$ as an IM proxy. Not only does it, on average, give high (low)

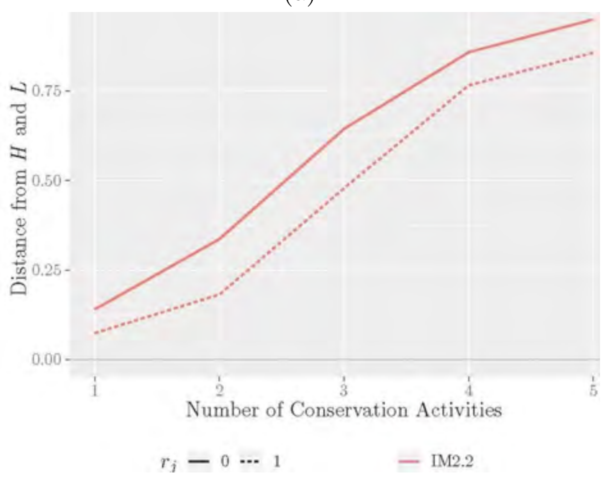
[7]Due to the rank-1 restriction on quantification matrix, the transformations on any other dimension is linearly dependent on the first dimension (principal component) [96].

values to farmers closer to H (L), but also penalizes the cases where the noise to signal ratio is high; whenever d_j^H and d_j^L are both high, $d_j^L - d_j^H$ will give relatively more precise information about intrinsic motivation than d_j^H and d_j^L individually. Figure 2.3a shows that $E_j(d_j^L - d_j^H|r_j)$ is strictly increasing in the total number of conservation activities for all $r_j \in \{0, 1\}$. Farmers who volunteer always score higher on this proxy than those who get paid, keeping N constant.

$d_j^L - d_j^H$ is just one out of many functions of $\{d_j^L, d_j^H\}$ which satisfy the properties mentioned in section 2.3. In fact, any function that is decreasing and increasing in d_j^H and d_j^L respectively, will satisfy those properties. One such function, $\text{logit}(d_j^L - d_j^H)$, is also used as a proxy for intrinsic motivation (see Figure 2.3b). Finally, Equation 2.2 is estimated using both $d_j^L - d_j^H$ and $\text{logit}(d_j^L - d_j^H)$ as proxies for intrinsic motivation which will be referred as IM-2.1 and IM-2.2 respectively in the rest of the paper.



(a)



(b)

Figure 2.3: (a) Distances d_j^H, d_j^L and IM2.1, and (b) IM2.2 as a function of farmers total number of conservation activities and their decision to get paid for them.

2.3.3 Method III

In method III, instead of directly predicting U_j^I as in subsection 2.3.1, I first estimate U_j and U_j^E and then find U_j^I using 2.1. Overall utility from conservation is approximated by performing Categorical Principal Component Analysis (CATPCA) on farmers' conservation practices data. Specifically, CATPCA is conducted on farmers' participation decision variables for soil, water, wildlife, pollinator and prairie conservation, and their total number of conservation activities. The participation decision for all conservation practices are treated nominally except N, which is considered an ordinal variable. Additionally, the quantification matrix for every variable is restricted to rank 1 [29]. Figure A.1.10 shows the optimally transformed variables plotted against their observed values. The plotted transformations are with respect to the first principal component.

Figure A.1.9 is the scree plot displaying the eigenvalues and variance accounted for (VAF) by the first three principal components. VAF for the first ('D1'), second ('D2') and third ('D3') principal components are 42%, 21.23% and 13.25% respectively. Component loadings in CATPCA, as in PCA, are correlations between the quantified variables and the principal components, and the sum of squared component loadings indicates the variance accounted for (VAF) by the principal components [90]. Table A.2.2 shows that D1 is dominated by N. D2, on the other hand, distinguishes farmers who participate in soil and water conservation from those who conserve pollinator, wildlife and prairie. A good proxy $\hat{U}_j$, like U_j, should be monotonic in N_j. Therefore, out of all the principal components, monotonic functions of D1 will be the most suited $\hat{U}_j$.

Next, $\hat{U}_j$ is regressed on the dummy variable for monetary payments r_j while controlling for other covariates as stated in Equation 2.4:

$$\hat{U}_j = \phi_0 + \phi_1 r_j + \zeta X_j + \nu_j. \tag{2.4}$$

Here, X_j include control variables such as farmer j's age, education, log of her farm size relative to the median farm size in her agricultural district and agricultural district fixed effects. ν_j is the random error and ϕ_1 captures how farmers' utility, $(\hat{U}_j)$, on average is affected by their propensity to accept payments for conservation. $\hat{U}_j - \phi_1 r_j$ then is the component of utility independent of payment. So, it is used as $\hat{U}_j^I$ and labeled IM-3.1.

Table A.2.4 summarizes the regression results where different functions of -D1 are used as $\hat{U}_j$. The second column uses -D1 as $\hat{U}_j$ while the first one is $E_i(\hat{U}_{ji}) = -\dfrac{D1}{N_j}$. In both specifications, getting paid, annually or onetime, significantly increases the utility from conservation.[8] The increase in total and average utility due to annual payment are 1.4 and 1.8 times more than the increase due to onetime payment. Since number of conservation practices is highly correlated with D1, the regression results also imply

[8]Note that since there are two types of payments available, annual and onetime, r_j and ϕ_1 are matrices of sizes $n \times 2$ and 2×1 respectively.

that getting paid on average increases the conservation activities of farmers. Higher education also increases the average utility from conservation but at a decreasing rate. Larger farms relative to the median farm size in their agricultural district significantly lowers the utility and consequently discourages conservation. This could be because the farms which are relatively large are likely to be more commercial in nature. Under such circumstances, taking land out of cropping to practice any type of conservation program may be too costly for the farmers.

I then subtract the effect of payments, annual and onetime, from the conservation utility to approximate intrinsic utility. The proxies for intrinsic utility created are: (i) $\hat{U}_j - \phi_1 r_j$ and (ii) $E_i(\hat{U}_{ji}) - \phi_1 r_j$ which are named as IM-3.1 and IM-3.2 respectively. Figure 2.4 shows how $\hat{U}_j^I$ on average varies with N_j and r_j $\forall\ j \in \{v, m\}$. As expected, the proxy based on total utility, $\hat{U}_j - \phi_1 r_j$, is strictly increasing in number of conservation activities. The intrinsic utility per conservation activity, on average, increases with N_j at a decreasing rate. But for both the proxies, the value is always lower for farmers of type m since their $\hat{U}_j^I = \hat{U}_j - \phi_1$ where $\phi_1 > 0$. As both of these proxies satisfy the properties mentioned in section 2.3, I use them to evaluate 2.2.[9]

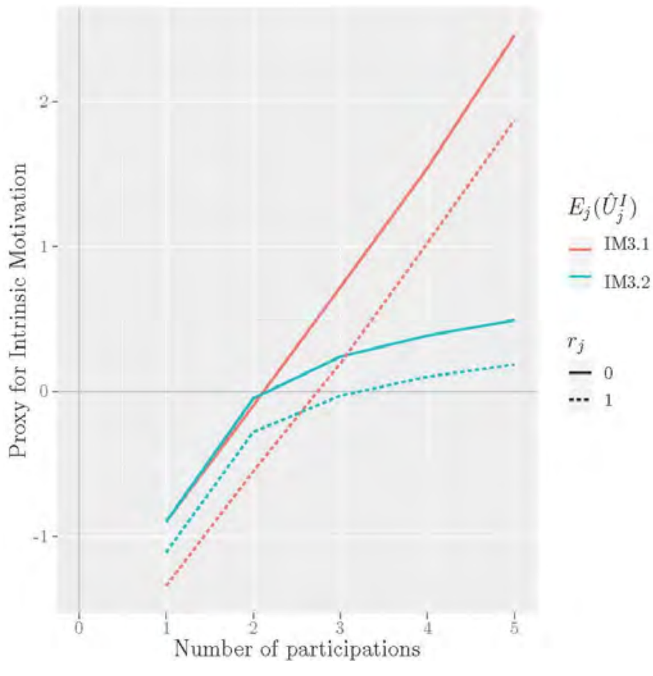


Figure 2.4: IM Proxies, $\hat{U}_j^I$, based on $\hat{U}_j = -D1$ and its average given the total number of conservation activities N_j and payment status r_j.

[9]Unlike Method I, my proxies here satisfy the properties in section 2.3. That is why, they are not weighted explicitly. Also, the variance of $\hat{U}_j^I$ here is almost constant across all N. Therefore, it is not a suitable weight for the proxy in question. The variation in $E_i(\hat{U}_{ji})$ however diminishes with increasing N and can be used to construct its weight.

2.4 Results

2.4.1 Summary Statistics

The mail survey, including a mail follow up for non-respondents, produced 599 (33.3%) returns. The mixed survey, also including a mail follow up for non-respondents, produced 75 (4.2%) online returns and 322 (17.9%) mail returns for a total of 397 (22.1%) returns. Out of all the survey returns, 698 are good returns that could be used for data analysis. Table A.2.5 summarizes the information on acres planted by the farmers, average yield of the types of crop they grow and whether they had livestock on their farm. 90% of the farmers report planting an average of 450 acres while renting out about 214 acres. Corn and soybean were grown by more than 60% of the farmers. 46% of the farmers have livestock on their farmland. Table A.2.6 summarizes the farm management practices of the farmers. It shows that non-invasive practices are more popular among farmers than their counterpart. Around 60% of the farmers report rotating crops, scouting, soil testing, using spring fertilizer and herbicide tolerant seeds. More than 50% do not use or have reduced tillage planting and report applying post as well as pre-emergent herbicide. Foliar insecticides, which is known to harm pollinators, is used by 20% of the farmers.

Table A.2.7 sums up the conservation practices of the farmers. 77.5% of the farmers practice some type of conservation activity. On average, farmers participate in environmental conservation activities more than animal conservation. The percentage of farmers participating in soil, water and prairie conservation are 58.7%, 34% and 18% respectively. On the contrary, only 31.6% and 16.4% have been involved in wildlife and pollinator conservation. Farmers with large farm area are less likely to participate in wildlife, pollinator and prairie conservation and more likely to conserve soil and water. Annual payment for conservation is the most common form of assistance. 40% of farmers who conserve, took annual payments while 11% took onetime payment. Getting paid is also associated with more conservation activities. Volunteers on average participate in 1.8 out of the 5 conservation activities. Conversely, those who get paid participate in 2.4 programs on average. This difference in average participation rate can be due to the significantly larger farms of paid participants. The farm size of volunteers and paid participants are 126.78 and 192.47 acres respectively.

The response rate is not statistically different across either the establishment or incentive types which lowers the possibility of selection bias in the survey data. 38.5% of the respondents agreed to participate in monarch conservation. Farmers who are involved in wildlife, pollinator and prairie conservation are more likely to participate in conserving monarch. 60% of these farmers agree to conserve monarch butterflies. But only 46% of those who do soil conservation agreed to establish monarch habitat. This relationship is however insignificant. Farmers decision to participate in monarch conservation is not significantly influenced by the proposed establishment methods. The percentage of farmers who agree to participate given establishment types A, B and C are 44.1%, 46.9% and 43.1% respectively. But, willingness to participate varies significantly with the amount of payment offered. Figure 2.5 shows that average willingness to establish milkweed decreases from 45.2% to 36.4% when payment level changes from 0 to $250. But further rise in the amount offered significantly increases average willingness to participate; at $1000 per acre onetime payment, the rate of participation is 59.6%. This result is akin to the findings of Gneezy *et al.* [48] and supports the possibility of crowding out of intrinsic motivation. Farmers agreeing to participate also indicated that they would enroll 2.45 acres of their non-cropland into the program on average.

Although establishment types do not significantly affect farmers' decision to participate in monarch conservation, yet farmers find some of them more manageable than others. They were asked to rate the establishment methods on their ease of compliance for various program requirements. These include location choice, pre-planting, planting, watering and mowing in the first and third years. Their options were *"not at all difficult"*, *"slightly difficult"*, *"somewhat difficult"*, *"Very difficult"* and *"Extremely difficult"*. Table A.2.8 shows the itemwise average difficulty level for each establishment type A, B and C. The average was calculated after the ratings were assigned values 1 to 5 in an increasing order of difficulty. I then test for equality of the average ratings using t-test. The establishment methods which are significantly easier to adapt to a program requirement are mentioned in the last column of the Table A.2.8. The location choice requirement is equally difficult across the establishment types A, B and C. But farmers significantly prefer establishment method C over B and B over A when pre-planting milkweeds. For every other requirement, they are indifferent between methods A and B. Nonetheless, method C is the easiest of all.

Farmers are also asked about how popular they think the program will be across the country −"If U.S. farmers in the monarch butterfly's migration path were invited to participate in this program, what percentage do you think would?". Their options included less than 1%, 1-4%, 5-9%, 10-24% and 25% and above. The results show that those who are optimistic about the program are significantly more likely to participate. Of those who anticipate more than 25% of U.S. farmers to participate, 78.6% agreed to do so themselves. On the other hand, only 30% of the farmers who think that overall participation rate will be less than 1% agree to participate in the program.

2.4.2 Regression results

I begin by examining how farmers' decisions to participate in monarch conservation program are affected by the amount of money they are offered, while not controlling for their intrinsic motivation. The participation decision variable is regressed on the dummy variables for dollar amount offered while controlling for other attributes of the program and farmers. These covariates include program attributes, farmers awareness about decline in monarch population, age, education, their farm size relative to the median farm in their agricultural district, and agricultural district fixed effect.

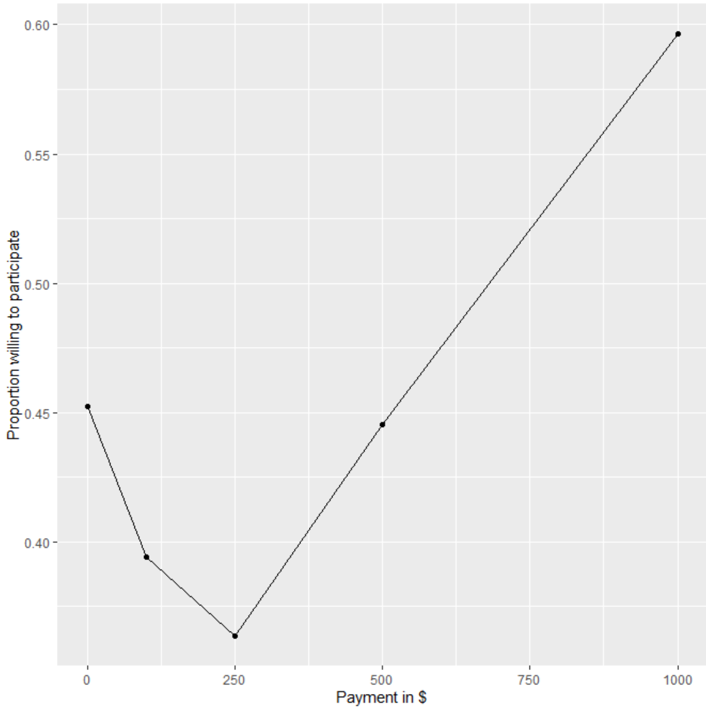


Figure 2.5: Percentage of respondents willing to participate in monarch conservation program as a function of payment offered.

Table A.2.9 summarizes the regression results using a probit model. To capture the effect of payment, following variables are used: First, the dummy variable 'Paid' which equals 1(0) when farmers are (not) offered money for participation. It captures the average effect that money has on willingness to participate. The second variable is '$value' which equals the per acre dollar amount proposed for participation. It takes values 0, 100, 250, 500 and 1000 respectively and will capture the marginal effect of money on willingness to participate. Figure 2.5 shows that introducing monetary payment causes a steep fall in the average level of participation in monarch conservation. But post introduction, the probability of participation becomes almost linear in payment. That is why, using these two variables instead of a separate dummy for each payment level, not only fairly represents the data but also helps gain degrees of freedom. The second specification has separate dummy variables for each level of conservation payment.

In the first specification, the probability of participation drops significantly by 42% when farmers get paid. Nonetheless, it recovers with further rise in payment level. An increase in payment by $1000 per acre will increase the chances of farmers conserving monarchs by 1%. This effect, although small, is highly significant. In specification 2, the participation rate declines with conservation payment until it is $250 per acre, where it reaches its minimum. Any further increase in the dollars offered raise the probability of participation. At $1000 per acre, the chance of participation is 33% higher than when no money is offered with a proability value of 11.7%. Farmers with at least 4 years of college are significantly more likely to participate than those who completed their high school at most. Relatively large farms are also less likely to participate, significantly so in case of specification 2.

Results in Table A.2.9 can be interpreted as crowding out of farmers willingness to participate by low conservation payments. Conversely, higher payment amount crowds in farmers' intrinsic utility which in turn enhances their average rate of participation. But the crowding out of motivation can also result from farmers engaging in socially desirable responding [51, 44]. Conservation activities, in general, are perceived to be altruistic and getting paid for doing the right thing is an unacceptable behaviour. Therefore, farmers may be over-reporting their intolerance towards conservation payments, especially when it is low, in order to appear virtuous. That is why, to test the crowding out hypothesis, their true taste for conservation is estimated from their current efforts to conserve. This variable is then interacted with the money offered for monarch conservation to see whether crowding out is indeed happening.

Figure A.1.11 shows the distribution of IM proxies which are re-scaled to lie in the interval [0,1]. The Shapiro-Wilk test of normality show that none of them follow normal distribution. The green and blue lines represent the mean and median of the IM proxy in question. Estimators 2.1, 2.2, 3.1 and 1.2 are all positively skewed. Accordingly, their median farmers are less intrinsically motivated than their average farmer. Proxy 1.1 on the other hand is negatively skewed. IM proxy 3.2 is also skewed to the left (-0.32) but its mode is less than its mean value. The bars in Figure A.1.12 show the average number of conservation activities of farmers in each quarter of IM proxies. The lines correspond to the secondary y-axis and represent the percentage of farmers who are paid participants in each quarter. All the proxies are monotonically increasing in the number of conservation activities. In other words, any farmer j who scores high on IM proxies (higher quartile) is more likely to have high N_j. Higher N_j will also incur more economic cost. Even intrinsically motivated farmers can find such cost inhibiting. That is why, in Figure A.1.12, the percentage of paid participants is not necessarily decreasing with IM scores except when the scores are based on IM1.1.[10] Conversely, when intrinsic motivation is approximated using IM1.2, the percentage of paid participants are statistically equal for all quarters based on a two-tailed proportion test. Alternatively, one-tailed test results illustrate significantly higher proportion of volunteers in its first quarter in comparison to its second quarter. Method II proxies have majority of volunteers in their second and third quarters. Their percentage are 40%, 61.8%, 66.7% and 51.3% for the first, second, third and fourth quarters respectively. In contrast, IM3.2 have volunteers mostly in its extreme quartiles. The percentage of voluntary participants in its first, second, third and fourth quarters are 67.5%, 41.67%, 47.5% and 60.5% respectively.

[10]This observation does not have a ceteris paribus assumption unlike the second property of IM in Section 2.3. Therefore, the two are not contradicting each other.

Next, I discuss the results from regressions where I use the variables defined in sections 2.3.1, 2.3.2 and 2.3.3 as IM proxies. The set of control variables will be the same as Table A.2.9 unless mentioned explicitly.

Table 2.1 summarizes the results from estimation of Equation 2.2 using IM proxies described in sections 2.3.1, 2.3.3 and 2.3.2 within a probit setup.[11] When discussing the results, the general emphasis will be on interpreting the signs and significance of the regression coefficients and not their absolute values, particularly for qualitative ordinal variables such as intrinsic motivation. The results are remarkably robust across all specifications. Farmers always prefer to volunteer for monarch conservation than getting paid, at any level of intrinsic motivation. Getting paid, on average, lowers the probability of establishing milkweed patches anywhere in the range of 43-47%. Yet offering more dollars to farmers with minimum intrinsic motivation significantly increases their average chance of growing milkweed, i.e. $\forall\, j$ whose $\hat{U}_j^I = 0$. A \$100 rise in conservation payment increases their likelihood of participation by 0.1% to 0.2% depending on underlying IM proxy.

	Intrinsic motivation proxy					
	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
Paid	-0.437^{**}	-0.463^{**}	-0.468^{**}	-0.471^{**}	-0.462^{**}	-0.477^{**}
	(0.192)	(0.195)	(0.195)	(0.195)	(0.194)	(0.195)
IM	0.791	1.336^{***}	0.945^{**}	0.835^{**}	0.909^{**}	0.881^{**}
	(0.725)	(0.491)	(0.416)	(0.330)	(0.438)	(0.358)
\$value	0.002^{**}	0.002^{***}	0.001^{***}	0.001^{***}	0.001^{**}	0.001^{**}
	(0.001)	(0.0005)	(0.0004)	(0.0003)	(0.0004)	(0.0005)
\$value×IM	-0.001	-0.002^{*}	$-0.001^{\dagger}$	$-0.001^{\dagger}$	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	-0.726	-0.746	-0.349	-0.266	-0.500	-0.611
	(0.616)	(0.517)	(0.494)	(0.493)	(0.500)	(0.506)
Observations	430	430	430	430	430	430
Log Likelihood	-267.056	-264.156	-265.247	-264.558	-265.408	-264.559
Akaike Inf. Crit.	608.112	602.312	604.494	603.115	604.816	603.118
Note:				†p<0.15; *p<0.1; **p<0.05; ***p<0.01		

Table 2.1: Summarizing regression results from IM proxies based on methods I, II and III.

Intrinsically motivated farmers, i.e. $\hat{U}_j^I > 0$, are, in general, more likely to participate in the proposed monarch conservation program. This is obvious since all the IM proxies have positive and significant coefficients, except in the case of specification 1.1 where it is positive but not significant at 10% level of significance. $\hat{U}_j^I$ interacts negatively

[11] See tables A.2.10, A.2.11 and A.2.12 for further details on regression results for proxies based on 2.3.1, 2.3.3 and 2.3.2 respectively.

with the amount of money offered signalling crowding out of intrinsic motivation. The coefficients of the interaction term range from -0.002 to -0.001. But, they are significant at 10% for IM1.2 only. The coefficients for interaction variable based on proxies 2.1 and 2.2 are significant at 15% level of significance. Although a less stringent criterion, using a threshold probability value of 15% to evaluate significance is not unreasonable here given the small sample size for farmers in group *H*.

Out of all the other control variables, education plays an important part in driving up the average probability of participation. Tables A.2.11, A.2.10 and A.2.12 show that farmers who have had at least four years of college are significantly more likely to establish milkweed patch than those who are at most high school drop outs. On the other hand, those who have had two years of college or any vocational degree have negative coefficients which implies a lower participation rate on average than those who are educated only up till high school. This difference, however, is insignificant across all specifications. Relatively large sized farms, on average, are less interested in establishing monarch habitat but this effect is insignificant. Farmers also do not have any preference for establishment types.

Table 2.1 also gives the AIC and log likelihood for each specification. Out of all, IM-1.2 has the lowest AIC and highest log likelihood value, indicating it to be the best fit for the data at hand. It should also be noted that the IM proxies which are non-decreasing and concave in N (here, 1.2, 2.2 and 3.2) fit the data marginally better than the others. Consequently, farmers' true intrinsic motivation may well be concave in their total number of conservation activities.

Figures 2.6 and 2.7 summarize the effect of money on the participation rate of farmers in different quartiles of IM1.2 and IM2.2 respectively. The key takeaways from Figure 2.6 are as follows. First, the probability of volunteering is strictly increasing in the intrinsic utility of farmers. On average, 70% of the farmers in the top quartile who were asked to voluntarily preserve monarch habitat are willing to do so. This number is 37.5% for those in the bottom quartile. Second, monetary rewards of \$100 and \$250 per acre cause the participation rate to decline among all the groups.[12] The fall is especially pronounced at \$250 per acre reward for farmers in the first and third quartiles. Their average participation declines from 37.5% and 58% to 20% and 28% respectively when payment offered increases from 0 to \$250 per acre. The decline however is only significant for the farmers is third quartile. Third, farmers in the bottom three quartiles are more willing to participate when payment rises further up from \$250 per acre. Fourth, the most intrinsically motivated group of farmers, i.e. those in top 25%, see a steady fall in their average probability of participation when the size of reward increases. Their

[12] At \$100 per acre, participation rate for all the groups is lower than when no money is offered, except for those in the bottom 25% of the distribution for the IM proxy. For them, the likelihood of participation increases marginally by 3.7% to become 41.2%. This rise, however, is highly insignificant with a p-value of 0.82.

participation rate decreases from 65% to 52% when their payment is increased from 0 to $1000 per acre. But, the difference in participation rate is not significantly different from zero. Consequently, the most motivated farmers can neither be lured in or out of their decision to conserve monarchs using money.

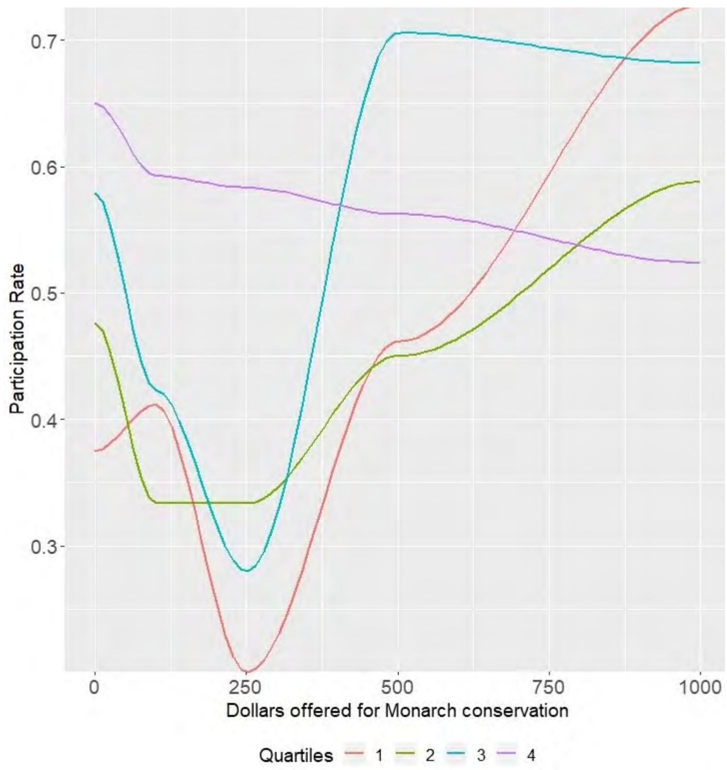


Figure 2.6: The effect of compensation level on the average probability of participation in monarch conservation for quartiles of IM-1.2.

Figure 2.7 is qualitatively similar to Figure 2.6 in many respects. Both these figures show that low levels of compensation ($100 and $250 per acre) deter participation of farmers in the bottom two quartiles, in addition to the third quartile in Figure 2.6. In Figure 2.7, the average probability of participation for farmers in the first and second quartiles decreases significantly from 46% to 13% and from 50% to 17% when their rewards are increased from zero to $100 per acre and $250 per acre respectively. But as soon as the conservation payment goes above $250 per acre, all these groups register a systematic increase in their participation rate. This suggest that farmers in the first and second quartiles of IM1.2 and IM2.2 as well as those in third quartile of IM1.2 may have a discontinuity in their utility functions at $0. Gneezy *et al.* [47] calls it the 'W-effect' of incentives. They state that performance is non-monotonic in the size of extrinsic incentives. When moving from no incentives to small ones, performance worsens, only to improve again with further increase in incentives. In one of their experiments, Gneezy *et al.* [48] find that pupils collecting donations for charity from private households perform better either when they receive a considerably large amount of money or when they do it for free. They mention several reasons for why incentives affect performance detrimentally, such as incomplete contracts and shift in perception from communal to exchange. I believe that introducing small conservation payments also change farmers' rationale behind participating in monarch conservation from pro-social to economic [78]. Small monetary rewards do not compensate them sufficiently and fail to gain their interest. The farmers may also perceive low levels of conservation payments as an insult to the time and effort they would have to put into establishing monarch habitat [48]. This significant negative effect is captured by the dummy variable on getting paid in Table 2.1. But as soon as the money offered becomes substantive ($500 or more), price effect overpowers any possible crowding out of intrinsic motivation and accordingly increases their average probability of participation. Due to this, the coefficient on the amount of conservation payment is significantly positive. The non-monotonic relationship between participation rate and conservation payments can also occur if farmers answered strategically [42, 18]. As per the survey, 74% of the farmers are aware of the declining monarch population, 51% of which report participating in Government conservation programs such as CRP, ACEP, CREP and WHIP to name a few. Many of these programs educate and also offer monetary rewards for pollinator habitat conservation, including monarchs, which in turn, update farmers about the urgency of the situation. Consequently, they can leverage their position as the potential providers of monarch habitats by strategically understating their willingness to participate in order to obtain higher compensation levels.

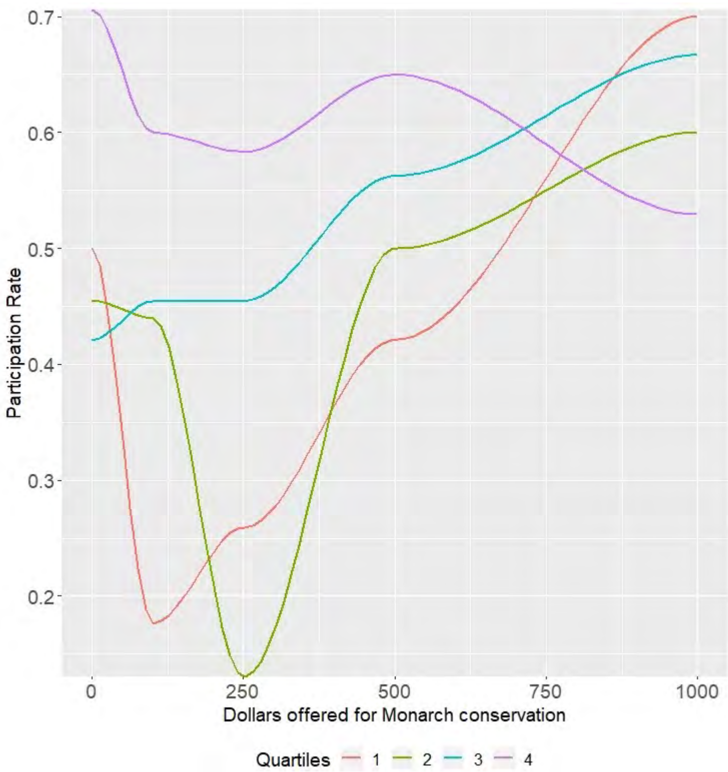


Figure 2.7: The effect of compensation level on the average probability of participation in monarch conservation for every quartile of IM-2.2.

The figures also show that the average participation rate of the most motivated groups, the fourth quartile, are less adversely affected by the low payment amount in comparison to the bottom two quartiles. In fact, the average participation rate for the third quartile of IM2.2 hardly decreases, even at low levels of compensation as seen in Figure 2.7. This may mean that their intrinsic utility from conservation is too high to be undermined by economic incentives in general. Consequently, in Table 2.1, even after netting out the effect of getting paid, high IM still improves the chances of conserving monarchs. But when conservation payments get very large, such as at $1000 per acre, the most motivated 25% experience decline in their average rate of participation. This deterioration is visible in both the figures and unfolds as the negative coefficient on the interaction terms between IM proxies and dollar value of conservation payment in Table 2.1. These effects become more apparent when an interaction between the IM proxy and dummy variable 'Paid' is included as an additional independent variable. Table 2.2 summarizes the results of probit regressions only for the best performing proxies, i.e. specifications 1.2, 2.2 and 3.2.[13] Adding the interaction term improves the log likelihood of every model. But, it causes AIC to deteriorate marginally for all specifications, except 2.1 and 2.2. The average probability of volunteering is no longer significantly higher for farmers with high IM values. Nonetheless, its coefficient is still positive, except for specification 2.1 where it equals -0.075 with a probability value of more than 0.9. Instead, part of the positive effect of intrinsic motivation is now captured by its interaction with the dummy variable for getting paid. This effect is significant at 5% for specifications 2.1 and 2.2 and at 21% for specification 3.1. Incorporating the new explanatory variable enhances the negative coefficient on the dummy variable for getting paid and its interaction with conservation payments. The coefficient on the dummy variable for getting paid almost doubles, although no longer significant at 10% for specifications 1.1 and 3.2. On the other hand, the interactions between the IM proxies and conservation payment are now significant at 5% level of significance for cases 1.2, 2.1 and 2.2.

In their paper, Gneezy *et al.* [48] state that insulting compensations are not necessarily small compensations. The top two quartiles together experience greater willingness to participate when conservation payments rise from $100 to $500 per acre. They may consider such modest rewards as a mere token of appreciation and consequently participate more. But when it further increases to $1000 per acres, their internal satisfaction from conservation reduces enough to decrease their willingness to conserve monarchs.[14] They may feel that their contribution will go unacknowledged when they are paid high sums of money for conserving monarchs.

2.5 Conclusion

In the past two decades, monarch butterflies have experienced a significant decline in their population, so much so, that they are being considered for listing under the Endangered Species Act of 1973. Researchers believe that the largescale reduction in milkweed,

[13] For further details, see Table A.2.13.

[14] Frey *et al.* [42] observe that Swiss residents when proposed hefty monetary compensations, in some cases higher than their median per month household income, to permit construction of a nuclear waste

	Intrinsic Motivation Proxy		
	(1.2)	(2.2)	(3.2)
IM	0.834	0.022	$0.730^{\dagger\dagger}$
	(0.722)	(0.500)	(0.546)
Paid	-0.876^{*}	-0.871^{***}	$-0.621^{\dagger}$
	(0.478)	(0.274)	(0.441)
$value	0.002^{***}	0.001^{***}	0.001^{**}
	(0.001)	(0.0003)	(0.001)
IM×Paid	0.881	1.317^{**}	0.268
	(0.932)	(0.634)	(0.735)
IM×$value	-0.002^{**}	-0.002^{**}	-0.001
	(0.001)	(0.001)	(0.001)
Constant	-0.487	0.026	-0.518
	(0.585)	(0.513)	(0.564)
Observations	430	430	430
Log Likelihood	-263.697	-262.387	-264.492
Akaike Inf. Crit.	603.395	600.774	604.984
Note:	††p<0.2; †p<0.15; *p<0.1; **p<0.05; ***p<0.01		

Table 2.2: Effect of interaction between IM proxies 1.2, 2.2 and 3.2, and getting paid on the probability of participation in the monarch conservation program.

which is monarch larvae's critical food source, since the introduction of herbicide tolerant crops plays a major role in aggravating the issue at hand. As a result, both the Government and other agencies in the U.S. are trying to develop plans which will encourage farmers to plant milkweed in their non-cropland. In this paper, I aim at finding out what factors encourage farmers to participate in monarch habitat restoration. Farmers in Iowa, Minnesota and Wisconsin are surveyed to evaluate their interest in hypothetical monarch conservation plans with predefined methods of establishing milkweed and monetary benefits. This paper especially focuses on how economic incentives affect farmers' willingness to participate, thus contributing further to the literature on motivation crowding theory. Farmers' intrinsic motivation is estimated using a variety of dimension reduction techniques on their actual conservation data.

The results suggest that monarch conservation will be more successful when farmers are asked to do it voluntarily. Incentivizing conservation economically will be a much more costly affair because of two reasons. First, intrinsic motivation is significantly crowded out at low levels of compensation, particularly for farmers who are inherently less inspired to undertake conservation activities. Second, while highly motivated farmers are not as averse to getting small conservation payments, they are, in any case, willing to participate, more so pro-bono. Thus my analysis implies that intrinsic motivation is non-monotonic in the size of monetary rewards, as conjectured by Gneezy *et al.* [48, 47].

repository for a short lived, low and mid-level radioactive waste on their community grounds, became less accepting of the idea than when no compensation was offered.

I also find that farmers are more likely to establish monarch habitats when assisted technically.

Another important takeaway is how to identify the right audience for a conservation program. Most of the studies testing public pilot projects are based on self reports from field experiments or interviews. This makes them susceptible to different types of biases such as social desirability bias and hypothetical bias thereby raising doubts on their external validity. In this paper, I estimate the respondents' chance of participating in monarch habitat conservation using data on their current conservation practices. Hence, my results are relatively robust and less vulnerable to biases mentioned above. Findings suggest that farmers who already volunteer in different types of conservation programs are more likely to establish milkweed patches on their non-cropland. Using actual conservation data to estimate intrinsic motivation also helps produce robust predictions about motivation crowding theory which in turn tell us about usefulness of financial instruments.

This paper unfolds multiple directions for future research. First and foremost, it is probable that respondents who get paid for conservation are inherently different from those who volunteer. Moreover, less intrinsically motivated respondents may as well over-report their willingness to volunteer in a hypothetical conservation program to look good. Such heterogeneity in the sample can bias in the estimates of intrinsic motivation. Therefore it becomes important to identify the psychological mechanisms behind people's decision to conserve. Additionally, it is important to outline the properties of intrinsic motivation. This paper offers a variety of intrinsic motivation proxies, all of which are characterized by the participants total number of conservation activities and their need for monetary assistance. Here, I have implicitly assumed that people give equal importance to every conservation practice to simplify the analysis. But the real world is more complex and it can be argued that different conservation activities are not comparable. Furthermore, the number of conservation programs people participate in is constrained by multiple factors in addition to their intrinsic motivation. For example, poor individuals with high environmental morale can find it difficult to transform their land into a ecosystem reserve without sufficient compensation. Under such circumstances, alternative definitions of intrinsic motivation can be recommended.

Chapter 3:

Valuing Monitoring Networks for New Pathogens: The Case of Soybean Rust

We estimate producers' value of information provided by alternative configurations of a monitoring network for an invasive wind-borne crop disease, soybean rust (*Phakopsora pachyrhizi*). Soybean producers use monitoring information to inform themselves about their risk of being infected and to make fungicide application decisions. The value of information is the expected gain in producers' profit associated with the use of the monitoring network, and we estimate this with a dynamic model of producer decision-making. We find that the value of the sentinel plot network increases with the number of sentinel plots, reaching a maximum at 400 sentinel plots. In addition, the optimal spatial arrangement of sentinel plots indicated by our model is substantially different from actual placements. Current sentinel plots are disproportionately placed in the Southern US where the risk of infection is high, but the amount of soybean is relatively small. Our estimates suggest more plots should be placed in the Corn Belt where the risk of an infection is lower, but where much more soybean is produced.

Keywords: monitoring network, value of information, Soybean Rust, sentinel plots.

3.1 Introduction

Soybean rust (*Phakopsora pachyrhizi*) arrived in the United States (US) in the late fall of 2004 [155], likely carried from South America by the winds of Hurricane Ivan [71]. Soybean rust is a fungal plant pathogen that cannot overwinter in temperate climates found in most of the US, but thrives throughout the year in the heat and humidity of the Gulf Coast, where sporulation in kudzu plants occur throughout the year. Rust spores are very likely blown into the continental interior of North America throughout the year, although many more spores are transported in the late summer and early fall because of the high availability of inoculum in the South. When weather conditions are suitable, rust spores deposited on soybean fields can take hold and cause substantial yield loss. Early indications of the potential damage from soybean rust were alarming. Based on experiences in Asia and South America where yield losses were high, Kuchler *et al.* [86] estimate that losses in the US could be as high as $7.1 billion per year, while Livingston *et al.* [92] find that losses in the first year of an infestation could run between

[1]This estimate is based on the yield loss reported in Wrather and Koenning [155] and the national average soybean prices for 2005-2008

$640 million and $1.3 billion. Damage has turned out to be less than feared: during 2005-2008, soybean rust caused an average loss of $4.14 million in the United States.[1]

Soybean rust is a recent example of a potentially catastrophic pathogen entering the US agricultural system and there are other plant diseases looming on the horizon. For example, the new races of wheat stem rust that recently emerged in Africa (Ug99) are of current concern, as are a number of pathogens that could affect corn and other crops [131]. In the US, the National Plant Diagnostic Network has been established to detect the arrival of new agricultural pathogens [144] and similar programs have been developed worldwide [107]. The arrival of soybean rust in the continental US prompted the creation of the Integrated Pest Management Pest Information Platform for Extension and Education (IpmPIPE, 2016), which includes a sentinel plot monitoring network. The sentinel plots in the monitoring network are areas of early maturing soybeans grown specifically to detect rust. The ipmPIPE also included a web-based information technology system that provides farmers with direct access to information on confirmed cases of rust, options for management, and forecasts and expert commentary on disease progression throughout the growing season. A successor project, the integrated Pest Information Platform for Extension and Education (iPiPE), has integrated parts of the soybean rust detection program in a larger effort designed to detect a wider array of pathogens.

A consortium of government agencies, agricultural trade organizations, and land-grant institutions provided the resources necessary to support the ipmPIPE. Such broad public and trade support may be justified by the public good aspects of the information provided [107] because farmers have limited incentives and ability to coordinate such efforts on their own. Still, since the adverse effect of soybean rust on US production has been less than anticipated, there have been questions about the value of the monitoring network as compared to its costs [91]. Particularly relevant is the work of Roberts *et al.* [127] which considers the case of the soybean rust network. They identify the value of the monitoring network as the benefit to farmers from making a better choice among three possible strategies, within a single planting season, due to a more accurate belief about the probability of infection. They find that the value of the network depends critically on the accuracy of the information and farmers' prior estimates of the probability of infection. For the most reasonable parameter values, and from an *ex ante* perspective, they find that the value of the program exceeded its costs. In recent years, investment in the soybean rust platform has been scaled back from more than 700 sentinel plots in 2007 to 87 sentinel plots in 2014 (USDA ipmPIPE restricted website).

In this paper, we focus on the value of a monitoring network such as the ipmPIPE program through modeling the optimal design of such a program. We build on the work of Roberts *et al.* [128, 127] and examine how the value of the network changes with additional sentinel plots included in the network and with re-allocations of the sentinel plots across space. The paper is organized in the following manner: We first characterize farmer behavior in the absence of a sentinel plot system. We then add a sentinel plot to the model so that farmers can use the sentinel plot to get a within-season signal about the presence of the pathogen. Next, we model multiple years of the network's operation since repeated observations of infection at a nearby sentinel plot allow farmers to learn about their vulnerability to infection even as they guard against it using a preventative fungicide. Then, we apply our model to US soybean acreage, starting with a set of county-specific prior probabilities of soybean rust infection. We introduce the possibility of a sentinel plot located in each county and compute the resulting value of information of each farm field-sentinel plot pair. We impose limits on the number of sentinel plots and use a spatial optimization model to locate them across US counties to maximize the expected gain in producers' profit associated with the use of the monitoring network. These results yield the value of information for different scales of the program. Our complete model allows us to determine values of a sentinel plot network, both contemporaneous and over time, and to find out how to get the most out of a budget-constrained program. Thus, our model can be used to assess the value of the current program as well as suggest changes to the current program that might improve efficiency.

3.2 Farmer Decision Model

In this section, we develop a farm-level decision model first in the absence of a sentinel plot network. We then introduce how a sentinel plot affects within-season decisions. Finally, we consider learning in a multi-season model.

3.2.1 Management without a Sentinel Plot Network

Following Roberts *et al.* [127], we assume a farmer faces two possible states of the world, represented by i: there is (i=1) or there is not (i=0) a soybean rust infection in her farm field. Also like Roberts *et al.*, we assume that a farmer can respond to the threat of a soybean rust infection with one of three management options: do nothing about the infection, scout and apply a curative fungicide as needed, or apply a prophylactic preventative fungicide. Let P be the price of soybeans, Y be the soybean yield when a farmer does not experience a soybean rust infection and C be the production cost exclusive of any effort to control soybean rust. Let λ_n be the proportion of yield lost in the event of an uncontrolled soybean rust infection. The returns to Do Nothing (N) strategy can then be written as $\pi_{N0} = PY - C$ without an infection and $\pi_{N1} = PY(1 - \lambda_n) - C$ with one.

One strategy is to scout and apply a curative fungicide as needed. We denote this strategy as the Curative strategy, labeled R. With this strategy, a farmer only applies fungicide when scouting reveals a soybean rust infection. It is costly to scout, and these costs must be incurred regardless of whether or not soybean rust is detected. If soybean rust is detected, then it is also costly to apply a curative fungicide. Furthermore, by the time soybean rust is detected, some damage to the crop will have already occurred. With $C_{sc} > 0$ equal to the cost of scouting, $C_c > 0$ equal to the cost of a curative fungicide, and λ_r equal to the proportion of yield loss with a curative fungicide in the event of an infection, the returns to this strategy can be written as $\pi_{R0} = PY - C - C_{sc}$ without an infection and $\pi_{R1} = PY(1 - \lambda_r) - C - C_{sc} - C_c$ with one. Note that, while a curative treatment will not give perfect control, losses from the Curative strategy will be less than losses from doing nothing in the event of an infection: $\lambda_r < \lambda_n$.

Choosing a Prophylactic Preventative strategy, labeled PP, instead of scouting or doing nothing about soybean rust reduces yield loss further because rust has little or no chance of damaging the crop. However, to be successful, a preventative fungicide must be applied before an infection occurs. Therefore, there is the possibility of applying a fungicide even when no infection would have occurred otherwise. Furthermore, the cost of a preventative fungicide treatment, C_p, is generally greater than the cost of a curative treatment: $C_p > C_c$. In the event of an infection, yield losses with a preventative treatment tend to be negligible [72], so we assume that λ_{PP} equals 0. Thus, the returns to the PP are written as $\pi_{PP0} = \pi_{PP1} = PY - C - C_p$ both with and without an infection.

The farmer's decision depends on the expected returns of the three strategies. If there is no chance of an infection, a farmer should clearly choose N. If an infection is sure to happen, returns depend on yield losses, fungicide costs, and scouting costs. Since a farmer chooses a strategy before knowing whether or not an infection will occur, the farmer's belief about the probability of experiencing an infection, ϕ^f, is crucial for choosing the best strategy. An important source of information for farmers is personal experience, which can be influenced by management choices. Choosing N or R allow them to observe, first-hand, any infection in their fields. These strategies provide them with an opportunity to better understand the underlying risk of infection. However, with a preventative fungicide, the farmers forgo this learning opportunity because treatment prohibits the emergence of the disease if the inoculum reaches the crop.

3.2.2 Within-Season Decision-making with a Sentinel Plot Network

In Roberts *et al.* [128, 127], the monitoring network sends one of two signals to a farmer: a high or low risk of infection. Farmers use the signal to update their beliefs about the probability of infection during the season. Further, the signal may be of low, medium, or high quality: the quality of the signal determines the degree to which farmers update their beliefs. Farmers can respond with one of three strategies: do nothing about soybean rust, scout and apply a curative fungicide in the event of an infection, or apply a prophylactic preventative fungicide.

Our approach differs from Roberts *et al.* [127] in three ways. First, in our model, we assume that the signal a grower receives is whether or not an infection has been confirmed at the relevant sentinel plot. This signal is more precisely defined than the high or low risk signal used by Roberts *et al.* (2009). Signal quality is expressed as the correlation between infection at the sentinel plot and infection in the farmer's field, and this correlation diminishes with distance. Second, we change the timing of decision making so that the farmer chooses her strategy based on her prior probability of infection before the signal arrives. Third, we add a fourth distinct management option: apply a conditional preventative fungicide mid-season if an infection is confirmed at the relevant sentinel plot during the season. We add this second type of preventative treatment to better characterize observed farmer behavior (Hershman). The Conditional Preventative (CP) strategy is defined as a strategy in which the farmer bases the decision of whether or not to apply preventative fungicide on whether or not the relevant sentinel plot becomes infected. This strategy is available only when the ipmPIPE and its monitoring network exists. If the ipmPIPE signals infection in the relevant sentinel plot, a fungicide is applied. Otherwise, no fungicide is applied. Note that the Conditional Preventative strategy is composed of the actions reflected in the Do Nothing (N) and Prophylactic Preventative (PP) strategies, but is different from these two strategies in the time of action. While the commitment to doing nothing or applying prophylactic preventative fungicide is made at the outset of the season with N and PP, with the Conditional Preventative strategy, the choice between the two actions is made once the signal from the sentinel plot has been received.

Given the ipmPIPE system, a farmer faces one of the following four situations: (1) there is an infection both in the farmer's field and in the sentinel plot, (2) there is no infection in either location, (3) there is an infection in the field but not in the sentinel plot and (4) there is no infection in the field and there is an infection in the sentinel plot. Let j represent the occurrence of soybean rust infection in the sentinel plot such that j equals 1 if an infection occurs, 0 otherwise. Let ϕ_{ij} be the subjective probability associated with each of the four scenarios described above. Thus, the expected return to the CP is $E(\pi_{CP}) = \phi_{11}\pi_{PP} + \phi_{10}\pi_{N1} + \phi_{01}\pi_{PP} + \phi_{00}\pi_{N0}$. Note that there is some probability (ϕ_{10}) that an infection will occur while the field goes unprotected because no infection is detected at the sentinel plot and N is chosen as the optimal strategy. Table 3.1 summarizes these scenarios and their corresponding probabilities.

		Field	
		Infection	No infection
Sentinel Plot	Infection	ϕ_{11}	ϕ_{01}
	No infection	ϕ_{10}	ϕ_{00}

Table 3.1: States of the world and their probability of occurrence when ipmPIPE is available.

The ipmPIPE provides farmers with information about the infection status of sentinel plots in the network. The extent to which this information helps farmers depends on correlation between the likelihoods of infection in the farmer's field and the sentinel plot. This correlation is expected to be decreasing in the distance between the field and the sentinel plot: a sentinel plot close to the farmer's field is likely to provide more useful information about the risk of infection in the field than one which is far away. However, nearness alone does not guarantee high correlation. Correlation also depends on the similarity in probability of infection at the field and sentinel plot. If a sentinel plot has conditions favoring the growth of soybean rust while a nearby field has hostile conditions for rust growth, then an infection in the sentinel plot does not necessarily imply a high risk of infection in the field. Thus, information from a sentinel plot is more useful to a farmer in revising beliefs if the sentinel plot faces a similar risk of infection as the farmer's field.

Let the subjective probability of infection in a sentinel plot be denoted by ϕ^s. We assume that beliefs about the probability of infection in any county are common knowledge. Let ρ be the correlation between the occurrence of infection in the farmer's field and sentinel plot. Then, each ϕ_{ij} can be expressed as a function of ρ, ϕ^f and ϕ^s in the following manner:[2]

$$\phi_{11} = \phi^f\phi^s + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{3.1}$$

$$\phi_{01} = (1-\phi^f)\phi^s - \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{3.2}$$

$$\phi_{10} = \phi^f(1-\phi^s) - \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} \tag{3.3}$$

$$\phi_{00} = (1-\phi^f)(1-\phi^s) + \rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}. \tag{3.4}$$

We know that $\phi_{11}, \phi_{10}, \phi_{01}, \phi_{00} \geq 0$. Therefore,

$$\rho \leq min\left\{\sqrt{\frac{\phi^s(1-\phi^f)}{\phi^f(1-\phi^s)}}, \sqrt{\frac{\phi^f(1-\phi^s)}{\phi^s(1-\phi^f)}}, 1\right\}. \tag{3.5}$$

We expect the correlation to be a decreasing function of distance *'d'* between the field and the sentinel plot. Although this sounds intuitive, there are no formal estimates of the relationship between correlation and distance. Therefore, we proceed by assuming the correlation to be inversely related to distance using the form $\frac{1}{1+e^{\theta d+\gamma}}$, where θ and γ are parameters which will be defined shortly. Also, ρ must be less than the R.H.S of equation 3.5 (UB). Hence, we scale the correlation function to ensure that ρ cannot be greater than UB:

$$\rho = \frac{UB}{1+e^{\theta d+\gamma}}. \tag{3.6}$$

We assume $\rho \in [0, 1]$ because a negative correlation would imply that a signal from the sentinel plot is systematically wrong. This would provide as much information as if the signal is systematically right. Instead, we assume that infections in distant sentinel plots

[2]Derivation for the following equations are shown in Appendix B.1

are uncorrelated with infections in the farm field.

3.2.3 Learning in a multi-season model

Recall that farmers' decision making capacity is expected to change with their personal experience because, as their experience increases, so does their understanding of their risks of rust infection. Their experience, in turn, is influenced by management choices. Choosing N or R allows them to observe any infection in their fields and therefore learn about their risk of infection. However, with a preventative fungicide, the farmers forgo this learning opportunity because treatment prohibits the emergence of the disease if the inoculum reaches the crop.

The ipmPIPE and its monitoring network provide an additional opportunity for a farmer to learn about the risk of infection in their field, irrespective of their chosen strategy. Because of the ipmPIPE system, the farmer knows if an infection happens in a sentinel plot. As a result, even if the farmer chooses to use a preventative treatment, information from the sentinel plot will help in refining the farmer's belief about the risk of infection in her field. The usefulness of this new information will depend on the correlation between the field and sentinel plot as demonstrated later. In essence, the ipmPIPE system provides many of the benefits of scouting without incurring any cost and allows a farmer to delay or completely forgo using a preventative treatment unless monitoring information suggests an infection has become likely (namely, if infections are identified at sentinel plots close to the farmer's field).

We now explain how our model captures a farmer's learning process about her probability of infection. We solve a multi-season decision problem wherein a farmer maximizes her expected utility of lifetime profits by choosing the best strategy for soybean rust control given her belief about the risk of infection and the availability of the ipmPIPE system. Let growing seasons be denoted by t. Suppose the probability that Y infections are observed in Z years in a region h follows a binomial distribution with parameter y_t^h for year t, where h can either denote the farmer's field or the sentinel plot. Also, suppose the farmer's prior belief about the probability of a rust infection in region h is characterized by the beta distribution with shape parameters α_t and β_t: $b(y_t^h; \alpha_t^h, \beta_t^h)$ with α_t^h as the number of times an infection is observed and β_t^h as the number of times a year passes without an infection in location h.[3] Then the farmer's state of the world in time period t can be defined by:

- $\{\alpha_t^f, \beta_t^f\}$ without ipmPIPE.
- $\{\alpha_t^f, \beta_t^f, \alpha_t^s, \beta_t^s\}$ with ipmPIPE.

Note that in the presence of the ipmPIPE and its monitoring network, the farmer has knowledge about the risk of infection in the sentinel plot in addition to the risk in her

[3] The beta distribution is an ideal way to describe a farmer's belief because it is constrained to the unit

farm. Using the beta distribution b to characterize the prior beliefs also implies that the posterior distribution $b(y^h_{t+1}; \alpha^h_{t+1}, \beta^h_{t+1})$ equals $b(y^h_{t+1}; \alpha^h_t + 1, \beta^h_t)$ when the farmer observes an infection and $b(y^h_{t+1}; \alpha^h_t, \beta^h_t + 1)$ when no infection is observed. The farmer's knowledge about her state of the world in $t+1$, or in other words, her learning depends on two factors: (i) the availability of the monitoring network and (ii) her optimal strategy for tackling the risk of infection in her field, given her beliefs.

- When the farmer chooses R or N, she gets first-hand information on whether or not her farm became infected. This is true even in the absence of monitoring network. Thus, α and β are updated in $t + 1$ when the strategy is either R or N, irrespective of the availability of monitoring network.
- When she chooses prophylactic measures, *PP*, she never gets to observe first-hand whether an infection would have happened in her field. But, with the monitoring network in place, she can refine her beliefs about the risk of infection in her field based on the information about infections in the relevant sentinel plot.
- Updating beliefs while choosing *CP* as the optimal strategy works in the same way as in case of *PP*. The only difference is that this strategy is available only when the monitoring network is present.

The knowledge about the state of the world in $t + 1$ translates into an updated belief about the probability of infection in the farm field. This is because the expectation of the beta distribution $b(y^h_t; \alpha^h_t, \beta^h_t)$ is

$$E(y^h_t) = \frac{\alpha^h_t}{\alpha^h_t + \beta^h_t} \tag{3.7}$$

and ϕ^f_t equals $E(y^f_t)$ and ϕ^s_t equals $E(y^s_t)$ by definition.[4] Hence, information on $\{\alpha^f_{t+1}, \beta^f_{t+1}\}$ and $\{\alpha^s_{t+1}, \beta^s_{t+1}\}$ in time period t characterizes the complete set of farmer's beliefs $\{\phi^f_{t+1}, \phi^s_{t+1}, \phi_{ijt+1},\ \rho_{t+1}\}$ for the next time period.

The farmer's expected return for the four possible strategies can thus be written as follows:

$$E(\pi_{N_t}) = \frac{\alpha^f_t}{\alpha^f_t + \beta^f_t}\pi_{N1} + \frac{\beta^f_t}{\alpha^f_t + \beta^f_t}\pi_{N0} \tag{3.8}$$

$$E(\pi_{R_t}) = \frac{\alpha^f_t}{\alpha^f_t + \beta^f_t}\pi_{R1} + \frac{\beta^f_t}{\alpha^f_t + \beta^f_t}\pi_{R0} \tag{3.9}$$

$$\begin{aligned} E(\pi_{PP_t}) &= \frac{\alpha^f_t}{\alpha^f_t + \beta^f_t}\pi_{PP1} + \frac{\beta^f_t}{\alpha^f_t + \beta^f_t}\pi_{PP0} \\ &= \pi_{PP0} \end{aligned} \tag{3.10}$$

$$E(\pi_{CP_t}) = \phi_{11t}\pi_{PP} + \phi_{10t}\pi_{N1} + \phi_{01t}\pi_{PP} + \phi_{00t}\pi_{N0}. \tag{3.11}$$

interval and is a conjugate prior of the binomial distribution.

[4] Since experiences with soybean rust are not the same across seasons, ϕ^f, ϕ^s, the ϕ_{ij}s and ρ are time dependent and are therefore denoted as ϕ^f_t, ϕ^s_t, ϕ_{ijt} and ρ_t respectively.

3.3 Dynamic Optimization

We now formalize our soybean rust problem using dynamic programming, where the state variables are the parameters of a beta distribution characterizing a farmer's belief about the probability of an infection. We use our model to learn how the optimal expected profit for a farmer improves in the presence of the ipmPIPE system. This increase in farmer's expected profit due to the availability of the monitoring network is defined as the value of monitoring. To compute the value of monitoring, we solve the dynamic programming problem both with and without the monitoring network.

3.3.1 Optimization without sentinel plot network

Without the monitoring network, the optimal value function for the farmer is:

$$V_t(\alpha_t^f, \beta_t^f) = \text{Max} \begin{cases} \mathrm{E}(\pi_{N_t}) + \delta \mathrm{E}(V_{t+1}(\alpha_{t+1}^f, \beta_{t+1}^f)), \\ \mathrm{E}(\pi_{R_t}) + \delta \mathrm{E}(V_{t+1}(\alpha_{t+1}^f, \beta_{t+1}^f)), \\ \mathrm{E}(\pi_{PP_t}) + \delta V_{t+1}(\alpha_t^f, \beta_t^f), \end{cases} \tag{3.12}$$

where

$$\mathrm{E}(V_{t+1}(\alpha_{t+1}^f, \beta_{t+1}^f)) = \phi_t^f V_{t+1}(\alpha_t^f + 1, \beta_t^f) + (1 - \phi_t^f) V_{t+1}(\alpha_t^f, \beta_t^f + 1) \tag{3.13}$$

is the expected value function when there are opportunities to learn and $\delta \in (0, 1)$ is the discount factor. The first argument on the right hand side of Equation 3.12 reflects a farmer's expected return from choosing the strategy N initially and the optimal strategy thereafter. Since choosing not to apply fungicide (doing nothing) lets the farmer observe whether or not an infection occurs, the farmer's optimal future strategies will depend on what is observed, and this yields the expected value function in Equation 3.13. With probability ϕ_t^f, the farmer will observe an infection and update their priors to $\{\alpha_t^f + 1, \beta_t^f\}$. With probability $1 - \phi_t^f$, the farmer will not observe an infection and then will update their priors to $\{\alpha_t^f, \beta_t^f + 1\}$. The second argument differs from the first only because the farmer chooses R initially. With this choice, the farmer still has the opportunity to learn, yielding the discounted expected value in 3.13 for the remainder of time. The third argument is different from the first two because the farmer chooses PP. This strategy does not provide any new information. Therefore, α^f and β^f remain the same in the next time period.

3.3.2 Optimization with sentinel plot network

Next, we find the farmer's optimal value function when the monitoring network is available. We assume that the network is available only for L years. The sentinel plots allow farmers to gather evidence about their risk of infection regardless of the management strategy they choose. In addition, the ipmPIPE and its sentinel plots provide a within-season signal upon which farmers can condition preventative treatments. Therefore, the optimal value function with the monitoring network (subscripted M) can be written as

$$V_{M_t}(\alpha_t^f, \beta_t^f, \alpha_t^s, \beta_t^s) = \text{Max} \begin{cases} \text{E}(\pi_{N_t}) + \delta \text{E}(V_{M_{t+1}}(\alpha_{t+1}^f, \beta_{t+1}^f, \alpha_{t+1}^s, \beta_{t+1}^s)), \\ \text{E}(\pi_{R_t}) + \delta \text{E}(V_{M_{t+1}}(\alpha_{t+1}^f, \beta_{t+1}^f, \alpha_{t+1}^s, \beta_{t+1}^s)), \\ \text{E}(\pi_{PP_t}) + \delta \text{E}(V'_{M_{t+1}}(\alpha_{t+1}^f, \beta_{t+1}^f, \alpha_{t+1}^s, \beta_{t+1}^s)), \\ \text{E}(\pi_{CP_t}) + \delta \text{E}(V'_{M_{t+1}}(\alpha_{t+1}^f, \beta_{t+1}^f, \alpha_{t+1}^s, \beta_{t+1}^s)), & t < L \\ V(\alpha_t^f, \beta_t^f), & t \geq L \end{cases} \tag{3.14}$$

where

$$\begin{aligned} \text{E}(V_{M_{t+1}}(\alpha_{t+1}^f, \beta_{t+1}^f, \alpha_{t+1}^s, \beta_{t+1}^s)) &= \phi_{11t} V_{M_{t+1}}(\alpha_t^f + 1, \beta_t^f, \alpha_t^s + 1, \beta_t^s) \\ &+ \phi_{00t} V_{M_{t+1}}(\alpha_t^f, \beta_t^f + 1, \alpha_t^s, \beta_t^s + 1) \\ &+ \phi_{10t} V_{M_{t+1}}(\alpha_t^f + 1, \beta_t^f, \alpha_t^s, \beta_t^s + 1) \\ &+ \phi_{01t} V_{M_{t+1}}(\alpha_t^f, \beta_t^f + 1, \alpha_t^s + 1, \beta_t^s) \end{aligned} \tag{3.15}$$

$$\begin{aligned} \text{E}(V'_{M_{t+1}}(\alpha_{t+1}^f, \beta_{t+1}^f, \alpha_{t+1}^s, \beta_{t+1}^s)) &= \phi_t^s \left[\frac{\phi_{11t}}{\phi_t^s} V_{M_{t+1}}(\alpha_t^f + 1, \beta_t^f, \alpha_t^s + 1, \beta_t^s) \right. \\ &+ \left. \frac{\phi_{01t}}{\phi_t^s} V_{M_{t+1}}(\alpha_t^f, \beta_t^f + 1, \alpha_t^s + 1, \beta_t^s) \right] \\ &+ (1 - \phi_t^s) \left[\frac{\phi_{00t}}{1 - \phi_t^s} V_{M_{t+1}}(\alpha_t^f, \beta_t^f + 1, \alpha_t^s, \beta_t^s + 1) \right. \\ &+ \left. \frac{\phi_{10t}}{1 - \phi_t^s} V_{M_{t+1}}(\alpha_t^f + 1, \beta_t^f, \alpha_t^s, \beta_t^s + 1) \right] \end{aligned} \tag{3.16}$$

are the expected value functions when the chosen strategy is either N or R (Equation 3.15), and when it is either CP or PP (Equation 3.16). If we look at equations 3.15 and 3.16 more closely, we find that they work out to be the same. However, they have different interpretations. At time period t when the farmer is choosing her optimal strategy, she knows that she will see the exact state she enters in $t+1$ if she chooses R or N. But, if she chooses CP or PP, she will not directly observe her state of the world in $t+1$ as these measures would prevent the emergence of infection in time period t altogether. Instead, now she predicts the state which she enters based on the information from sentinel plots. As a result, while Equation 3.15 captures the actual expected value of the farmer, Equation 3.16 gives only an expectation of the expected value. The sentinel plots also provide a within-season signal to condition preventative treatments i.e. use CP, yielding an additional opportunity in the maximum of Equation 3.14 that did not appear in Equation 3.12. This strategy has the contemporaneous expected return of $E(\pi_{CP_t})$. Once the sentinel plots are no longer available (when $t \geq L$), the farmer's problem becomes identical to the problem in equations 3.12 and 3.13.

It should also be noted that in our model farmers can use monitoring network free of cost.[5] If the monitoring network usage was costly then equations 3.15 and 3.16 would not have been the same. In this case, the farmer would have used the information from monitoring network only when choosing PP or CP. Had she chosen N or R as her optimal strategy, she would have already been able to perfectly update her belief based on her direct observation of whether infection occurred in her field or not. Here paying to get information from the sentinel plot would not have honed her belief about her risk of infection any better. Hence, if the monitoring network was costly, the expected future value in 3.15 would have been a function of only $\{\alpha^f_{t+1}, \beta^f_{t+1}\}$ and not of $\{\alpha^s_{t+1}, \beta^s_{t+1}\}$. On the other hand, expected future value in 3.16 would still have been a function of $\{\alpha^f_{t+1}, \beta^f_{t+1}, \alpha^s_{t+1}, \beta^s_{t+1}\}$.

3.4 Farm Model Implications

Having these two models of decision making allow us to calculate the value to farmers of the monitoring provided by sentinel plots. Recall that the value of monitoring at any time is the difference in the expected value of managing soybean rust with and without information. Therefore, at $t = 0$, the value of information is: $V_{M_0}(\alpha_0, \beta_0) - V_0(\alpha_0, \beta_0)$. Solving equations 3.12-3.13 or 3.14-3.16 requires information on: expected yield losses in the event of an infection both with and without curative treatments (λ_r, λ_n); expected soybean yields and prices (Y and P); scouting and fungicide costs (C_{sc}, C_c, and C_p); a farmer's initial beliefs about the probability of a soybean rust infection $\{\alpha_0, \beta_0\}$; the number of years that the monitoring network is available (L); the distance and correlation between infection at the farmer's field and sentinel plot (d, ρ); and the discount factor (δ). L is assumed to be 4 years. The values used for λ_r, λ_n and δ were adopted from Roberts *et al.* [128, 127] and are summarized in Table B.1.1.[6] Their values for C_{sc}, C_c and C_p were for 2005 and are adjusted to account for inflation. We use distance between counties as an estimate of d. These distances are calculated based on the Haversine formula. As a result, in our model, a field and sentinel plot in the same county have distance $d = 0$ between them. We assume ρ to be very high (≈ 1) at distances less than 80-100 mi. We also assume that the correlation drops after this point and that its value approaches 0 when the distance reaches 200 miles. Given these assumptions, and if we suppose correlation to be 0.01 at 150 mi, then parameters θ and γ from Equation 3.6 take on the values 0.12 and -13.4, respectively.

[5] This is consistent with the actual model of ipmPIPE.

[6] Additional tables and figures for this chapter are in and named after appendix B and are duly numbered.

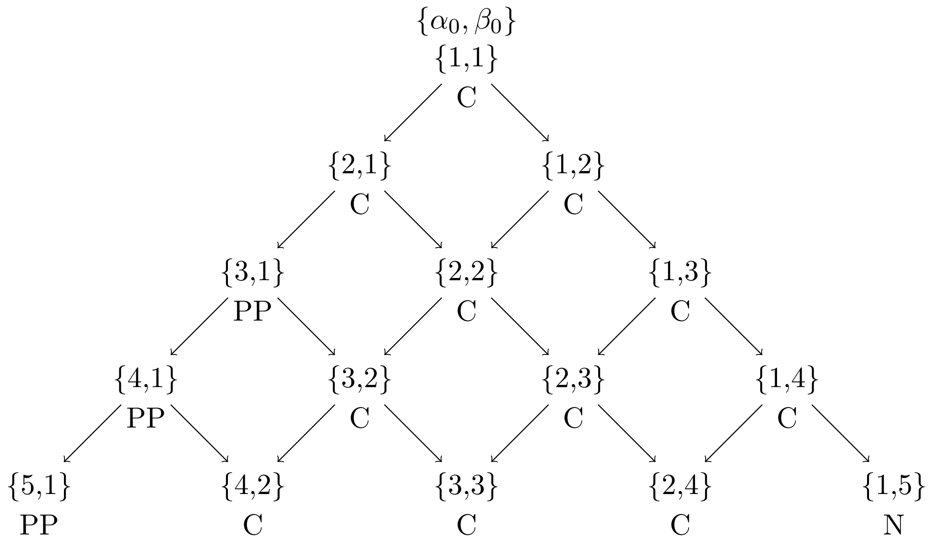
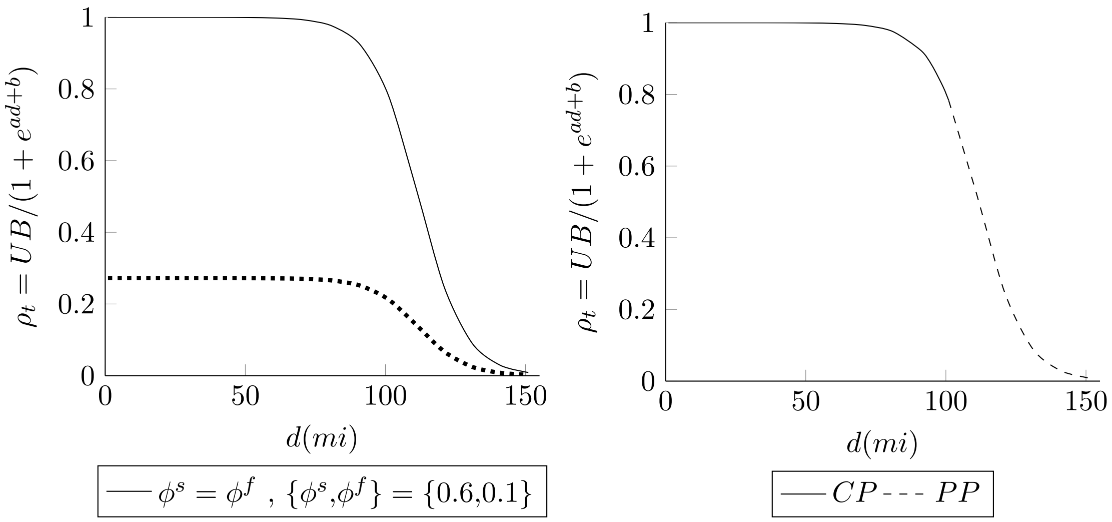


Figure 3.1: Optimal management decisions for a farmer with no monitoring network.

Before turning to the primary questions of interest regarding the overall value of the sentinel plot monitoring network, we explore some of the basic implications of the farmer decision model. We solve the dynamic programming model for an infinite planning horizon using the parameter values in Table B.1.1 and an assumed yield of 37 bushels/acre. Figure 3.1 shows the first four years of optimal strategies at each time period given the farmer's beliefs about the probability of infection in absence of monitoring network. The vertices of the decision tree are labeled according to the state variables α_t^f and β_t^f. At the first vertex, α_0^f and β_0^f are both equal to 1 and the beta distribution is equivalent to the uniform distribution. Branches to the left represent observed infections and branches to the right are years without infections. This figure shows that the first node where it is optimal for farmers to choose PP is with a prior belief described by $\{3, 1\}$. This corresponds to the first 2 observations being infections. Alternatively, if a farmer starts with a uniform prior and sees no infections for the next four years $\{1, 5\}$, then it is optimal to choose N. For the majority of states in this example, it is optimal for the farmer to choose R.

Figure 3.2b shows how optimal strategies change in presence of the monitoring network. Suppose beliefs about the probability of rust infection are 0.58 for both the farmer's field and the sentinel plot ($\alpha_0^f = \alpha_0^s = 5.8$, $\beta_0^f = \beta_0^s = 4.2$). According to our model, when the sentinel plots are unavailable, the farmer chooses PP as the optimal strategy. When the sentinel plot is present in the same county as the farmer's field, i.e., when $d = 0$ and $\rho = 1$, she will use the information from the monitoring network since it will tell her whether her field will get infected or not. Therefore, whenever the sentinel plot is in the same county as the farmer's field, CP is the farmer's optimal strategy. However, the monitoring network becomes less informative with an increase in distance d. This is because an increase in d decreases ρ and consequently the value of information from the monitoring network. Up to $d < 100$ mi, the model predicts that the farmer will continue with CP. Once the sentinel plot is placed at a distance $d \geq 100$ mi, our model predicts that the farmer will no longer rely on the information from monitoring network and will revert back to the optimal strategy without a monitoring network, PP.



(a)

(b)

Figure 3.2: (a) Correlation function ρ and (b) optimal management decisions for a farmer with monitoring network when $\phi^f = \phi^s$.

It should be noted that the threshold where the optimal strategy shifts from CP to a different strategy depends on the shape of the underlying correlation function. Suppose there is no monitoring available and the belief about the probability of infection in the field (ϕ^f) is equal to 0.1. With this low probability, the farmer's optimal strategy is N. Now, suppose there is a nearby sentinel plot with ϕ^s equal to 0.6. The farmer's field and the sentinel plot are characterized by very different beliefs and therefore have a very low correlation ρ irrespective of the distance d between them (Figure 3.2a). As a result, even if the sentinel plot is placed close to the field, say at d=5 miles, the sentinel plot does not provide valuable information about the likelihood of infection in the farmer's field. Hence, the farmer chooses to continue doing nothing about soybean rust even in the presence of the monitoring network.

3.5 Spatial Optimization Model

We implement our farm-level model assuming a representative farmer in each county, using the 2014 soybean acreage, yield and price data. The source of the data is USDA-NASS. While the yield and acreage are at the county level, the data on soybean prices is at the state level. We use county level estimates of the probability of a soybean rust infection from Bekkerman *et al.* [17] to represent prior beliefs about the likelihood of a rust infection for the representative farmer in each county. To transform these estimates into valid parameters for a beta distribution, we calculate α_0^h and β_0^h such that

$\frac{\alpha_0^h}{\alpha_0^h+\beta_0^h}$ equals the probability estimate and $\alpha_0^h+\beta_0^h=10$, which ensures that all priors are based on the same number of years of potential rust infection observations prior to our evaluation.[7] Basic summary statistics for these variables are reported in Table B.1.2.

Given yields, prices, costs and farmers' initial beliefs, we now allocate a fixed number of sentinel plots to the counties and maximize the value of information generated from them. We formulate this spatial optimization problem using a linear integer programming model. Let $m \in \{1, 2...M\}$ represent the counties with farm fields and $n \in \{1, 2...N\}$ represent counties with sentinel plots. Then, the value of monitoring for county m when sentinel plot is placed in county n is: $v_{mn} = V_M(\alpha_0^m, \beta_0^m, \alpha_0^n, \beta_0^n, L) - V(\alpha_0^m, \beta_0^m)$. We define z_n as a binary decision variable equal to 1 if county n is selected for a sentinel plot. Let x_{mn} be a binary decision variable equal to 1 if sentinel plot in county n is assigned to farmers in county m. Then the spatial optimization problem can be stated as:

$$\begin{aligned}\max_{x_{mn}, z_n \in \{0,1\}} \quad & \Sigma_{m=1}^{M} \Sigma_{n=1}^{N} v_{mn} x_{mn} \\ & x_{mn} \leq z_n \quad \forall m, n \\ & \Sigma_{n=1}^{N} x_{mn} = 1 \quad \forall m \\ & \Sigma_{n=1}^{N} z_n \leq B.\end{aligned}$$

The objective is to maximize the value of information across all M counties. The first constraint stipulates that the assignment of county n to county m can take place (i.e., $x_{mn} = 1$) only if county n is selected for a sentinel plot (i.e., $z_n = 1$). The second constraint requires that each county m with a farm field is assigned one county n with a sentinel plot. The third constraint stipulates that the number of counties selected for sentinel plots must be less than the budget, B, where the budget is defined as the number of sentinel plots in the system.

We identify 1360 counties for which we have all information required: initial beliefs, yields, acreage and prices for 2014. Hence, in our spatial optimization problem, $M = N = 1360$. We compute the value of monitoring for each pair of these 1360 counties, leading to 1360×1360 values of monitoring (v_{mn}). An exact solution to this spatial optimization problem is then solved using the branch and bound algorithm in GAMS with the Cplex solver [45].

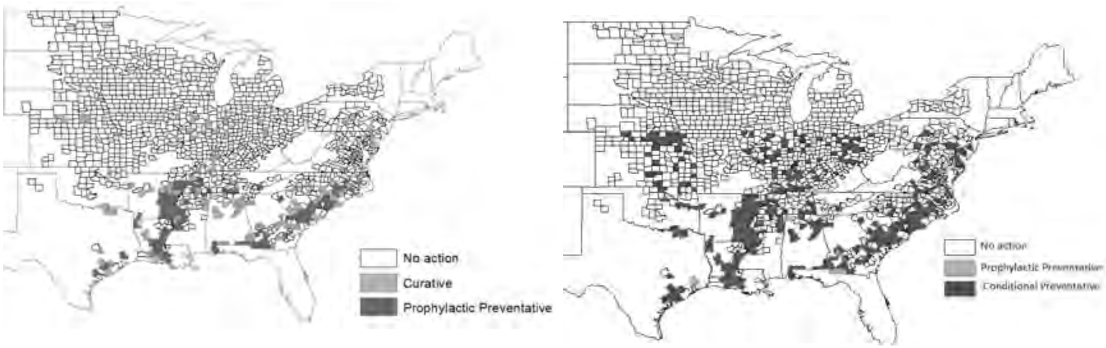
3.6 Results

In our model, factors such as the probability of infection and production (yield and acreage) play important roles in the farmer's decision making process. Based on these

[7] Soybean Rust was first diagnosed in 2004 while our study is for 2014. Therefore there are 10 obser-

factors, we can predict the optimal strategy of farmers both with and without the monitoring network. We start by studying figures B.1.1 and B.1.2 in order to understand the intuition behind our solution before moving on to discussing the solution itself.[8]

When there is no monitoring network, farmers can choose either N, R, or PP to control rust infection in their fields. Doing nothing about soybean rust (N) leads to maximum profit in the case of no infection and minimum profit when infection occurs. Therefore, it will be the most suitable for farmers who have a very low chance of rust infection. On the other hand, PP will be optimal for those who expect severe losses from soybean rust. This will generally be true for farmers with a very high risk of infection. But, when a farmer faces a high probability of infection along with comparatively low soybean production, she may find that the costs of a preventative strategy outweigh its benefits. In this case, she will choose scouting (R) over preventative strategy (PP).



(a) (b)

Figure 3.3: Optimal strategies when (a) monitoring network is unavailable and (b) post-spatial optimization when B=80.

The figures also show that while soybean is mostly grown in the Midwest, the probability of rust infection in the Midwest is typically less than 0.02. Hence, in Figure 3.3a, we see that most of the northern states choose N. The Southeast is generally characterized by a high probability of rust infection owing to its proximity to Gulf of Mexico. That is why we observe the choice of prophylactic treatment PP in Arkansas, Mississippi, Louisiana and coastal counties. The counties in which farmers choose PP have an average probability of infection of 0.68 and average yield of 47.21 bu/acre in comparison to the corresponding values being 0.01 and 45.72 bu/acre for those who optimally do nothing about rust infection. The counties where R is the optimal strategy are more moderate in nature. They face an average risk of infection (0.39) which is high enough to abandon N as optimal strategy but lower than the risk in areas using prophylactic treatment. These regions also have a comparatively low soybean yield (44.04 bu/acre). This lower yield, when combined with moderate/high risk of infection, does not make the expected loss from a rust infection high enough to justify the extra cost of prophylactic treatment over scouting. For example, farmers in counties in northern Texas and

vations on rust infection available to each farmer about her field.

[8]The areas in white are not included in this study.

Alabama and in northwestern South Carolina, which have comparatively higher probability of infection but low soybean yield, choose R.

The optimal strategy of a farmer is also a function of the soybean price she faces. Based on our model, the expected loss is increasing in soybean prices. Hence, everything else constant, we can say that higher prices would lead to more preventative treatment. In our data, soybean prices are at the state level and do not have much variability as shown in Table B.1.2. Hence, when interpreting the results we shall discuss the effect of factors such as production and risk of infection on farmer's value of monitoring and optimal strategy rather than the effect of soybean prices.

Figures B.1.1 and B.1.2 also help us understand how the value of monitoring varies spatially across counties. The value of monitoring should increase with the total expected loss from rust infection and therefore should be increasing in the probability of rust infection, soybean acreage and soybean yield. Based on this logic, we can make a few easy predictions: first, the value of monitoring will not be large for states like Minnesota, Wisconsin, North and South Dakota owing to their very low likelihoods of infection. Second, the value of monitoring is not expected to be very big for most of the counties in southeastern states such as Alabama, Georgia and South Carolina which are characterized by low soybean acreage and yield. Third, the counties on the border of Arkansas and Mississippi should be among the major beneficiaries of the monitoring network because they have high soybean acreages and yields and they experience a high likelihood of rust infection.

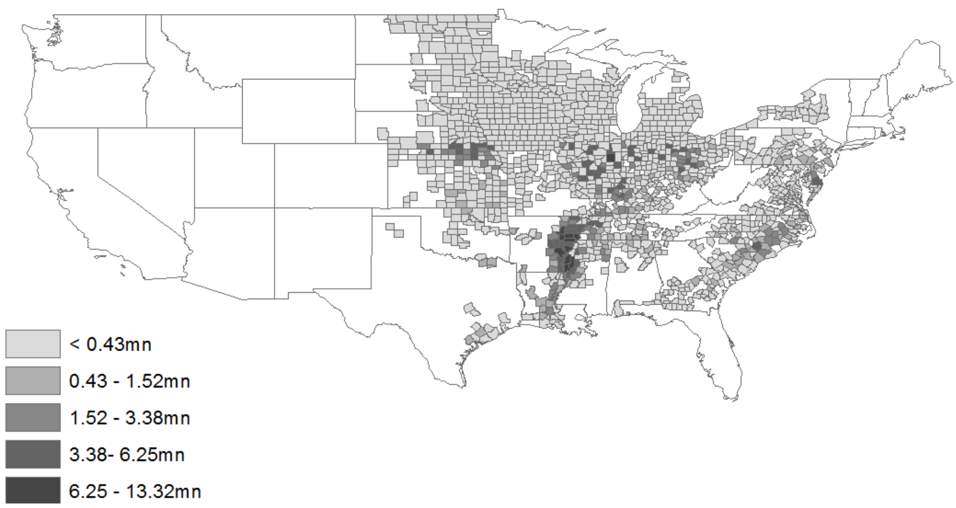


Figure 3.4: Value of monitoring (in millions) when each county has a sentinel plot.

Figure 3.4 shows that our results match our intuition. It displays the dollar value of monitoring when all 1360 counties have sentinel plots. While the northern and southeastern states do not gain much from monitoring in general, Arkansas and Mississippi are among those who profit the most from sentinel plots. In this region, most of the counties have soybean grown on at least 60,000 acres while facing a risk of rust infection of at least 50%. It also shows a high value of monitoring for many counties in the states of Nebraska, Illinois, Indiana and Ohio that do not have high probabilities of rust infection but are characterized by high enough yields and acreage to result in large expected losses from soybean rust.

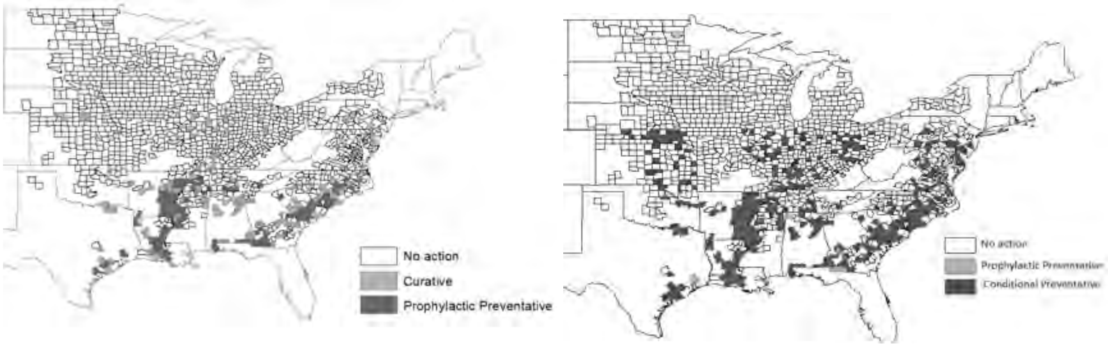
Figure 3.4 can be interpreted as the solution to the spatial optimization problem when B takes its maximum value, i.e., when $B = 1360$. We now study the optimal placement of sentinel plots when the number of sentinel plots, B, decreases. Figure 3.5 shows how sentinel plots are allocated among counties to maximize the total value of monitoring in US when B equals 185 and 80 respectively. When B equals 185 (Figure 3.5a), sentinel plots are allocated to states with the highest values of monitoring such as Arkansas, Mississippi, Illinois and Ohio, as well as to southern Nebraska. Reducing the number of plots to 80 causes Nebraska, Ohio and Illinois to lose many of its sentinel plots as seen in Figure 3.5b. These regions, although characterized by high soybean production both in terms of yield and acreage, have comparatively low risks of rust infection (at most 25%). As a result, the expected loss from soybean rust is not big enough in these places to preserve the initial allotment of sentinel plots when B shrinks from 185 to 80. A big cluster of sentinel plots remains in Arkansas and Mississippi.

(a) (b)

Figure 3.5: Optimal sentinel plot locations when (a) B=185 and (b) B=80.

Figure 3.6 shows optimal management strategies with and without the monitoring network. There are two facts which are clear when we compare figures 3.6a and 3.6b. First, in the presence of a spatially optimized monitoring network, most farmers switch to using CP as their optimal strategy.[9] Second, with optimally allocated sentinel plots farmers no longer use Curative treatments. Both these results clearly point out how the farmers prefer to use signals from monitoring network in order to make better pest management decisions. However, while these points do depict the benefits of a monitoring network to farmers, it should be remembered that these benefits also incorporate the gains from optimal spatial allocation of sentinel plots.

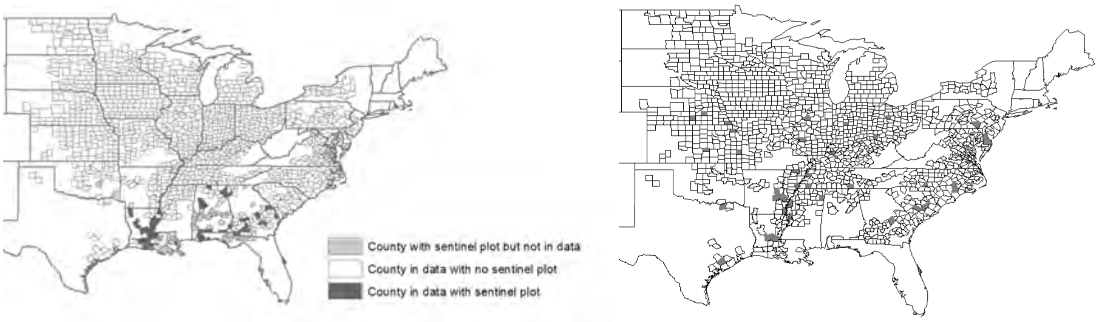
[9]In fact, when B=1360 farmers only use Conditional Preventative.



(a) (b)

Figure 3.6: Optimal strategies when (a) monitoring network is unavailable and (b) post-spatial optimization when B=80.

Figure 3.6 also demonstrates that counties derive spillover gains from nearby sentinel plots. Figure 3.5b shows that 3 sentinel plots have been allocated close to the state of Nebraska. In Figure 3.6b, we see that this area has 10 more counties that use information from nearby sentinel plots, i.e. choose CP, to tackle rust infection. Therefore, not only are the sentinel plots beneficial to the farmers in their own county, but they also benefit those in other counties. Such gains are feasible in our model when the correlation between the farm field and sentinel plot is high. This requires meeting two conditions (Figure 3.2b): i) the counties should have similar probabilities of infection, and ii) the counties should not be separated by a very large distance. Here, both these conditions are met. Not only are these counties neighbours, but they also have similar beliefs about the probability of infection as seen in Figure B.1.1. This makes the correlation sufficiently high for these counties to value the information from a nearby sentinel plot.



(a) (b)

Figure 3.7: Counties with (a) sentinel plots in 2014 and (b) sentinel plots post-spatial optimization when B=39.

We now compare how the spatial optimization of sentinel plots could improve the total value of monitoring relative to actual placement of plots in 2014. Figure 3.7a maps counties with sentinel plots in 2014. There were 75 counties in total that contained sentinel plots. Out of these 75, only 39 counties are included in our data. Hence, for the comparison to be appropriate, we solve the spatial optimization problem with $B = 39$. Figure 3.7b shows the optimal placement of sentinel plots. Examining these two figures simultaneously shows that the actual sentinel plot locations in 2014 do not match the counties identified as optimal in our model. As predicted earlier, the optimal solution would place sentinel plots mostly in northern Mississippi, Tennessee, Arkansas, Louisiana, and in the Midwest region. However, sentinel plots in 2014 were mostly clustered in the southern part of the Southeast region (e.g., Alabama, Florida, Georgia, and Louisiana). Although this region has a moderate to high risk of infection, it has very low soybean production and is also far away from the main production areas. As a result, the quality of signal of infection coming from these far off sentinel plots are relatively poor for the main soybean producing areas. This results in a total value of monitoring of \$0.07 bn at maximum. On the other hand, the optimal value of monitoring post-spatial optimization is \$0.566 bn, which is 808.6% of the actual value captured. It should be noted that, while our findings show the inefficient allocation of sentinel plots, we do not wish to emphasize the magnitude of the estimated loss from this inefficiency. This is because our estimates of the total value of monitoring are dependant on the assumptions about the correlation function. A higher correlation across longer distances would lead to a higher value of monitoring based on the sentinel plot locations in 2014.

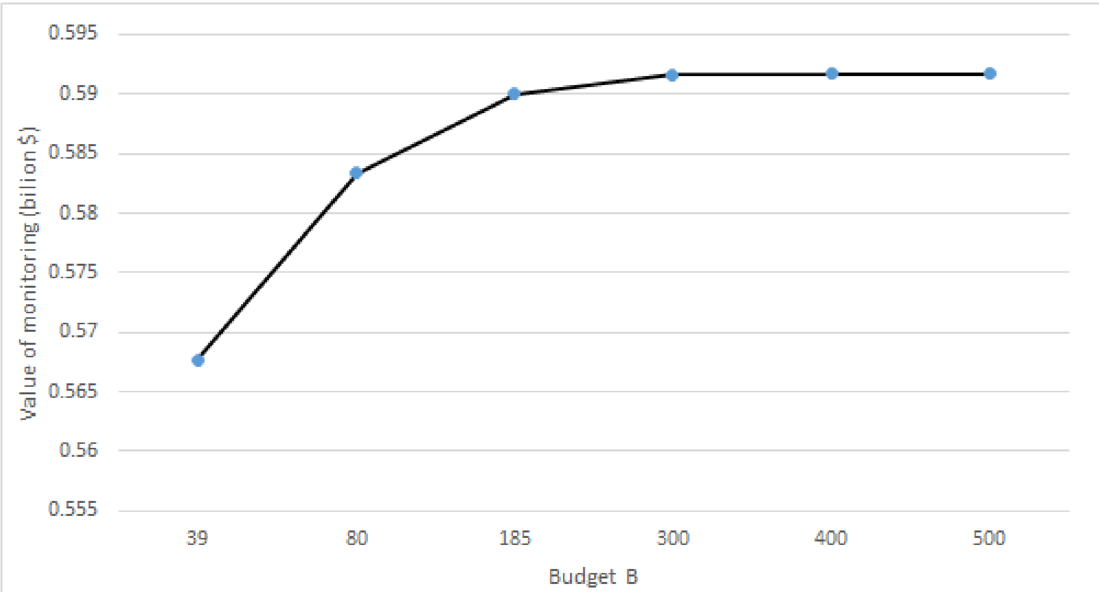


Figure 3.8: Value of monitoring by number of sentinel plots (budget).

Figure 3.8 shows how the optimized value of monitoring varies with budget B. When B is 80, the total value of monitoring is \$0.58 bn. This value keeps increasing at a decreasing rate until B is 400 where it attains a maximum of \$0.59 bn. No additional value is gained from sentinel plots beyond 400, i.e., the marginal benefit of sentinel plots becomes zero. Because we do not explore the costs associated with expansion and maintenance of the sentinel plots network, we cannot determine the optimal number of plots. Nonetheless, it is clear that there is no additional benefit in having B greater than 400, assuming the sentinel plots have also been efficiently allocated to the counties.

3.7 Conclusion

Potentially catastrophic plant pathogens threaten to enter the US every year. In order to prevent damage from these pathogens, the government invests in sentinel plot monitoring networks to help predict and therefore control the risk of plant diseases. However, the net benefit of such monitoring networks is greatly debated because they are costly to maintain and, for many pathogens such as soybean rust, very few incidents of infection have been reported throughout the US mainland. In this paper, we develop a dynamic model of farmer decisions regarding pest management and use it to estimate the value of the sentinel plot monitoring network and to optimize the spatial arrangement of sentinel plots. The paper uses the case of soybean rust and hence estimates the value of ipmPIPE, the sentinel plot monitoring network for soybean rust.

Farmers benefit from the information provided by the monitoring network since it helps them manage pests better within growing season and also refines their beliefs about their risk of infection. They value information more when their expected loss (average soybean production times the risk of infection) from infection is high given the quality of information. The quality of information from a sentinel plot decreases when the sentinel plot is placed far away from a farmer's field. As a result, when the sentinel plots are placed in the Deep South while soybean is mostly being grown in the north, the total welfare from monitoring will be lower than when they are placed further north. Current sentinel plots are disproportionately placed in the Southern US where the risk of infection is high, but the amount of soybean is relatively low. Our estimates suggest that more plots should be placed in the Corn Belt where the risk of an infection is lower, but where much more soybean is produced. Such a modification in sentinel plots arrangement could have increased the value of monitoring in 2014 by 808%. However, it is important to note that there are other potentially valuable uses of monitoring plots that we did not include in our model of farmers' decisions. For example, monitoring plots in the South not only inform local soybean growers about management decisions but also inform aerobiology forecasts of soybean rust spread in the corn-belt, which inform growers about when and where to scout for soybean rust infections. Further, monitoring plots for soybean rust in the South also provide information about the presence of other soybean diseases. Further work is needed to account for these additional uses and update the valuation and optimal location of monitoring plots.

Our results also show that, given our assumptions, no more than 400 sentinel plots are needed to maximize the value of ipmPIPE. There have been multiple years, especially in the initial years of its life, when ipmPIPE had many more than 400 sentinel plots. In recent years, there have been decreases in the number of sentinel plots. However, most of the sentinel plots still remain in the South. This fact points us to a useful direction of research– estimating the optimal number of sentinel plots in ipmPIPE. With data on the cost of maintenance and establishment of sentinel plots for each county, one can estimate the marginal cost of monitoring. The optimal number of sentinel plots can then be estimated by equating the marginal cost of monitoring to its marginal value.

Before we answer this question, it is important to better understand the spatial autocorrelation of infections since the marginal benefit curve is derived based on assumptions about the correlation function. We assume that the correlation of risk of infection is sigmoidal in the distance between sentinel plot and farmer's field and that it becomes zero at 150 mi from the farmer's field. These assumptions are not yet corroborated. Any change in these assumptions can cause significant changes in the value of monitoring and therefore in the suggested arrangement of sentinel plots. This leaves scope for improvement.

Chapter 4:

Will Farmers Adopt Remote Sensing for Soybean Aphid Management? An Economic Perspective.

This study examines the economic suitability of Unmanned Aerial Vehicles (UAVs) for scouting soybean aphids based on a plant-level spatiotemporal bioeconomic model of infestation. We find that UAV based scouting, although imprecise, generates optimal profit equivalent to manual scouting. But its greater tendency to detect false positives relative to manual scouting can also trigger frequent unnecessary treatments and dramatically reduce farmers' profits. Yet, UAV's commercial viability depends more on reducing its operating cost than improving its precision, once it has a tally threshold of 250 soybean aphids per plant.

Keywords: Integrated Pest Management, Unmanned Aerial Vehicle (UAV), Soybean aphid, agent-based model.

4.1 Introduction

Globally, soybean is the largest source of animal protein and second largest source of vegetable oil. The U.S. is the world's largest soybean producer with an average value of over $40 billion produced on more than 80 million acres annually. The most important insect management challenge facing U.S. soybean farmers is the soybean aphid, an invasive pest that was first discovered in the U.S. in 2000. In Minnesota, soybean aphid has reduced the plant height, pod numbers, seed size and quality, and yield [115]. Soybean aphid is also a carrier of a number of plant viruses [28]. The growth rate of soybean aphid is known to be rapid in controlled environment i.e. under ideal conditions where population growth is not perturbed by the effects of weather or natural enemies.[1] Yield losses from this relatively new U.S. pest can reach 40% without management [122, 81]. To meet this management challenge, farmers have primarily planted soybean seed treated with neonicotinoid insecticides and applied foliar pyrethroid insecticides [67, 80].

The profitability of aphid management using neonicotinoid seed treatments has been questioned because the efficacy of the systemic neonicotinoid insecticides declines over

Support for this research was provided by the University of Minnesota's MnDRIVE Transdisciplinary Research initiative and the Minnesota Agricultural Experiment Station project MIN-14-134.

[1] It is reported to double in 1.5 d in controlled environment [101]. However, the estimates of the soybean population growth from the field is known to be less than the theoretical intrinsic rate of growth [25]

time limiting its control of later season aphids [54]. Soybean aphid populations from Minnesota and Iowa exhibited resistance ratios up by 40-fold for pyrethroids (i.e. bifenthrin and lambda-cyhalothrin) between 2015 and 2016. The reports of pyrethroids failing to control soybean aphid in the field were noted from Minnesota, Iowa, North Dakota, and South Dakota [80]. This questionable profitability combined with evidence of the negative impacts of neonicotinoid insecticides on pollinators has the U.S. Environmental Protection Agency debating restrictions on their use in soybean production [50, 34, 49]. Regardless of the outcome of this debate, there is increasing pressure for soybean farmers to use more integrated pest management practices, which would include an increased reliance on foliar insecticides.

The profitable and sustainable use of foliar insecticides requires application only when the aphid population has reached an economically significant level. To help soybean farmers better time their foliar applications, entomologists have developed and refined manual scouting protocols that randomly sample soybean plants in order to estimate the severity of aphid infestation. An insecticide application is triggered if the average number of aphids per plant exceeds the profit maximizing threshold (250 aphids per plant is the current recommendation) [123, 62]. This protocol is repeated throughout the growing season as long as aphids are active. It is labor intensive and not very precise given the small number of plants that can be economically sampled, which helps to explain why growers have relied more on planting insecticide treated seed.

Recent research using unmanned aerial vehicles (UAVs) shows that they can detect the presence or absence of aphids above the recommended threshold in 1.5×2.5 meter grid cells within a soybean field. While these detection rates are not without error, this discovery raises the prospect of replacing manual scouting with UAV based scouting protocols that observe an entire field rather than a small sample from it. The technology is currently not yet advanced enough to provide precise estimates of the pests or diseases responsible for observed stress or the severity of infestation, advances in navigation and imaging hardware and software are making remote sensed scouting costs competitive with conventional scouting costs. As the costs of obtaining remotely sensed scouting information declines, an interesting question that arises is how precise does remote scouting information have to be before it is more profitable to use than conventional scouting? In this paper, we intend to explore this questions.

This objectives will be accomplished through the development of an agent based pest management model and the application of this model to soybean aphid management. In this model, individual plants are the agents. Each plant is initially and randomly endowed with an insect population (or is insect free). Each day the insect population on a plant grows and then a portion of it migrates to other plants. At the end of the growing season, cumulative aphid days is calculated and used to determine a plant's yield loss. Within this biological framework, various decision rules can be constructed for simulating treatments that reduce the plant's insect population. For this paper, two decision rules are considered. Both the decision rules are based on a weekly scouting regimen with whole field insecticide applications triggered when scouting detects an insect infestation that is severe enough to reduce expected profitability without the application. The first decision rule is manual scouting where an insecticide application is triggered when the average number of insects on a random sample of plants exceeds the treatment threshold. The second decision rule is UAV scouting where an insecticide application is triggered when the proportion of plants with a detectable (with error) aphid population exceeds the treatment proportion. For each of these decision rules, the expected net return to management inclusive of scouting costs, average number of insecticide applications, and average total insecticide use are calculated using Monte Carlo methods.

The key results of the paper are as follows. First, when we assume that the per hectare cost of scouting is the same for UAV and manual scouting, and that insecticide treatment following UAV scouting is triggered only if soybean aphid density is at least 250 aphids per plant, then the optimal expected profit based on UAV scouting is \$785.18 per hectare. Conversely, the optimal expected profit from manual scouting is \$785.3 per hectare. With manual scouting, treatment is suggested in third and fourth weeks since the arrival of soybean aphids once the optimal threshold of 275 aphids per plant is reached. In case of remote sensing, the field should be treated in the fourth week once 45% of its plants are infested with at least 250 aphids per plant. Second, while the two methods have very similar optimal profits, manual scouting is still a better option because it estimates infestation level more precisely. It is significantly better for farmers who do not experience soybean aphid infestation frequently or those who prefer treating their fields early in the season. Third, expected profit from UAV scouting does not vary much with its probabilities of type I and type II errors. When it goes from providing perfect information to no information at all, the optimal expected profit falls from \$785.18 to \$778.15 per hectare. Fourth, increase in cost of UAV based scouting causes a linear decline in its profitability with no affect on its economic threshold. Pederson *et al.* [118] find UAV based weed scouting to be 80% as costly as the conventional scouting. Under such circumstances, remote sensing for surveying soybean aphid will result in an expected profit of \$796.02 per hectare which is a 1.4% more than the expected profit from manual scouting. If the cost of operating UAVs became negligible, then its expected profitability will rise to \$840 per hectare.

The paper is structured in the following way: Section 4.2 reviews the literature on soybean aphids scouting methods, the use of UAVs in agriculture and spatial bio-economic models which are used in studying plant diseases. Section 4.3 explains our spatial model of soybean aphid infestation using cellular automation with agent based models. We specify the biological processes ranging such as reproduction, migration, wing induction and death, which together form state transition rules in subsections 4.3.1.1, 4.3.1.2 and

4.3.1.3. Section 4.3.2 describes the pest management strategies available to the farmers and their profit maximization problem within the biological framework mentioned above. Our model is parametrized in section 4.4. We then elaborate how we simulated our model in section 4.5 and our results in section 2.4. Finally, section 4.7 concludes.

4.2 Literature Review

Integrated Pest Management (IPM) for soybean aphid uses repeated sampling throughout most of the growing season. In Minnesota, it is recommended to sample 20-30 random soybean plants from all over the field beginning late vegetative stage through reproductive stage R5 [79]. Farmers are also recommended to treat their entire soybean field with foliar insecticides when aphid population exceeds 250 aphids per plant with over 80% of plants infested and population is increasing [62]. This Economic Threshold (ET) of 250 soybean aphids per plant is estimated based on the theory of economic injury level (EIL) [122]. EIL is defined as the lowest pest population density that causes economic damage. ET is defined as the injury equivalency of a pest population corresponding to the latest possible date a given control tactic could be implemented to prevent injury from causing economic damage [119]. Ragsdale *et al.* [122] estimate EIL ($\approx$ 674 aphids per plant) and ET ($\approx$ 273 aphids/plant) between R1 and R5 (assuming 4-d lag) based on 19 yield loss experiments conducted over a period of 3 years in six states. If left untreated, soybean aphid herbivory can cause yield loss exceeding 40%.

The ET of 250 aphids per plant is based on a static model and lacks a dynamically optimal decision guide when multiple treatments are needed [157]. Its recommendations can be suboptimal because of two reasons. First, since this model is static, it assumes that the marginal benefit of all treatments are independent and equal. But treating crops today affects how beneficial subsequent treatments will be later in the season. Decision making based on the static model, therefore, may overestimate the value added by a treatment and consequently recommend more than needed. Second, static models do not incorporate the cost of increasing pest resistance to insecticides on optimal control of a pest. Regev *et al.* [66] show that when making treatment decisions, farmers only take into consideration the monetary cost of insecticides and not the increased user costs. User costs are increased future costs of controlling the pests as a result of the decision to apply chemicals today. If user costs are accounted for, it would result in the ET increasing during the course of the growing season rather than remaining fixed at a particular level as usually presented in entomological literature.

The recommended ET assumes that aphid population will reach EIL if not treated. But there are multiple biotic and abiotic factors such as deteriorating host plant quality, crowding, natural enemies and extreme weather which can prevent the escalation of population to EIL [148, 76, 88]. In fact, soybean aphids are known to have a density dependent population [32]. Costamagna *et al.* [24] proposes an exponential growth model with intrinsic growth rate decreasing linearly with time. They attribute the decreasing rate of growth to deteriorating plant phenology with time. Matis *et al.* [100] on the other hand suggest a population model where density dependence is introduced via death rate being dependent on cumulative population.

Cantangui *et al.* [22] introduced stage-specific EILs and ETs for R2, R4 and R5 soybean development stages using the law of diminishing increment regression model and symmetric bell shaped and logistic growth models. They perform caged experiments in 2003 and 2004 for their study, which suggested leveling off of the yield loss with increasing soybean aphid numbers. This motivates using the regression model based on the law of diminishing increment. They argue that stage specific EILs can give the farmers enough lead time to decide which developmental stage of the plant is the most suitable time for them to treat for soybean aphid. For infestation starting at V5, the stage specific EILs are 3.5, 74.6 and 212.1 soybean aphids/plant at R2, R4 and R5 respectively assuming a field with yield potential of 3700 kg/ha, soybean market value of $0.29/kg and control cost of $24.7/ha. This study has been criticized because it bases EILs on caged plants which prevents the access of natural enemies to aphids and any aphid movement from or to the host plant [114]. As a result, the estimated EILs are likely lower than optimum, which can cause significant overtreatment facilitating the development of an insecticide resistant soybean aphid population.

Hodgson *et al.* [63] developed a binomial sequential sampling plan using field collected data in Minnesota from 2001 to 2003 and computer simulations of sampling effort. This binomial sequential sampling plan underlies speed scouting. Speed scouting is based on the mathematical relationship between proportion of plants infested, aphid density per plant and ET of 250 aphids/plant. Instead of taking whole plant counts, speed scouting introduces a tally threshold (40 aphids/plant) to declare plants as infested or not infested. As a result, as few as 11 plants are needed to make decision about treatment. Hodgson *et al.* [61] test the validity of speed scouting using commercial fields in Minnesota and replicated small field trials in Iowa, Michigan, Minnesota and Wisconsin. They conclude that 79% of the time speed scouting resulted in same recommendation as whole plant counting based on ET of 250 aphids/plant. However, speed scouting is a conservative plan and hence it consistently recommends insecticide use before ET is reached using whole plant count.

Alves *et al.* [6] find that UAVs can also detect stress due to soybean aphids. In their experiment, soybean aphid decreased the near-infrared reflectance (NIR) and Normalized Difference Vegetation Index (NDVI) reflectance, which provide measure of plant health. However, similar changes in NIR and NDVI can be detected due to other types of stressors such as soybean cyst nematode and soybean sudden death syndrome [57, 55, 15]. Marston *et al.* [98] examine combination of features (wavelengths) of a UAV which will improve its accuracy in detecting aphid pressure. They also use machine learning, in particular linear Support Vector Machine (SVM) models, to generate an actionable information from remote sensing data. They find that all wavelength combinations trained with SVM models were able to classify the aphid pressure above and below ET of 250 aphids per plant with over 80% accuracy. However, the models were more accurate in identifying cases with aphids below ET than above ET.

Treatment strategies are based on imperfect monitoring of pest population. Fackler *et al.* [38, 39] discuss how such observational uncertainty can be addressed using an extended Partially Observable Markov Decision Processes (POMDP). In context of pest management, POMDP approach replaces the imperfect information on pest population with a probability distribution and actions are taken based on this distribution. These type of models are also capable of handling structural uncertainties which occur due to the imperfect knowledge of the underlying theory [154]. Haight and Polasky [52] model the problem of monitoring and treating a site for an invasive species using POMDP with infestation level as a state variable. They find that quality and cost of monitoring changes the optimal management strategy. With costless and perfect monitoring, expected costs are 20–30% lower across the range of belief states relative to the expected costs without monitoring.

There is a large literature on control of pests by their natural enemies which suppresses pest population growth and has potential to mitigate pest control costs and crop yield loss. Zhang and Swinton [157] exploit this predator-prey relationship in modelling managerial choices. They develop an intra-seasonal dynamic bioeconomic optimization model for insecticide based pest management that takes into account both the biological control effect of natural enemies on the pest population and the effects of pesticide on the level of natural pest control supplied. Thus, they introduce insecticide decisions using a natural enemy adjusted economic threshold. However, they do not take into account the relative voracity of natural enemies. Bahlai *et al.* [13] develop a mechanistic dynamic tritrophic population and phenology model for soybean aphid, incorporating environmental cues, host plant cues and natural enemy dynamics. To standardize the impact of different predators, they introduce the concept of Natural Enemy Unit (NEU) which is defined as the number of natural enemies that kill 100 individual prey in 24 hours. Their results suggest that natural enemy abundance and host plant phenology are the most important factors affecting soybean aphid population. One of the major drawback of this study is that it is deterministic and non-spatial. Hallett *et al.* [53] estimate a dynamic action threshold (DAT) in presence of natural enemies by introducing NEUs in the population growth model. In their model, an insecticide application was triggered only if natural enemy numbers were insufficient to suppress pest populations. DAT also provided equivalent yields to the conventional action threshold during field experiments and reduced the average number of pesticide applications.

Spatial bioeconomic models have become popular recently in studying plant diseases and infestation. Sanchirico *et al.* [133] investigate how ignoring spatial distribution of bioeconomic models can lead to suboptimal results. Space can be incorporated in bioeconomic disease models using state transition probabilities based on location [124] and partial differential equations [65]. These models however assume spatial heterogeneity to be exogenous [136]. But this may not be true. The spread of pests or disease to a new plant mostly depend on whether its neighboring plants are infested or not. Agent based models (ABM) allow spatial heterogeneity to be determined endogenously. These models study population of heterogeneous autonomous agents. These agents make decisions based on simple rules by interacting with other agents and their environment [146]. Atallah *et al.* [11] use ABM to investigate the profitability of different disease control strategies for grapevine leafroll disease. They model grapevine leafroll disease at plant level using cellular automata and find that spatial strategies targeting immediate neighbors of symptomatic vines are better than nonspatial strategies.

4.3 Bioeconomic Model

This section explains the bioeconomic model used in this paper. It has two subparts. The first part is about modeling infestation of soybean aphids in a field during a single growing season. To do this, we use cellular automaton (CA) model with ABM. Miller [106] define CA as discrete spatio-temporal dynamic systems based on local rules. They are characterized by regular lattice of cells, all of which have a finite number of states. There are rules to govern the state transition of cells in every time period; a cell transitions into a new state based on its current state and the states of its neighbors according to the transition rules [146]. In our model, soybean field is represented by a grid and its cells are occupied by soybean plant agents. The state of a cell or a soybean plant is defined by the number and composition of its resident soybean aphid as well as its location. State transition of a cell is determined by pertinent biological processes of soybean aphids such as their reproduction, migration, wing induction and death. The second part of our model deals with profit maximizing pest management strategies for the farmer. We analyze pest management decisions based on manual scouting and aerial surveillance (UAVs).

4.3.1 Population Dynamics of Soybean Aphids

The soybean field is a two dimensional grid with M_1 rows and M_2 columns and consequently there are a total of $M_1 \times M_2$ cells. Grid cells are of equal size and are denoted by i where $i \in \{1, 2....M_1M_2\}$.[2] Each cell has a soybean plant situated at its center. Suppose the soybean plants in the field are homogeneous at any given time.

Consider a single soybean growing season. Soybean aphids enter the field at time $t=0$. They are characterized by their morph j and age group l and so, are referred as

[2]Although each cell is characterized by 2-D (row, column), here we identify them by a single uniquely defined index i which is a linear combination of its row and column number.

soybean aphids $\{j, l\}$. Their morphs include alate (winged, w) and aptera (wingless, $\tilde{w}$) i.e. $j = \{w, \tilde{w}\}$. The age groups are nymph ($\tilde{a}$) and adult (a) i.e. $l = \{a, \tilde{a}\}$. We consider a soybean aphid to be nymph only on the day it is born; it becomes an adult next day onwards.[3] Let the population of aphids $\{j, l\}$ at time t in cell i be denoted by N_t^{ijl}. Therefore, total number of aphids in cell i, labeled N_t^i, equals $\sum_j \sum_l N_t^{ijl}$.[4] The state of a cell i at time t is given by vector $S(i|t)$ such that

$$S(i|t) = (N_t^{i\tilde{w}\tilde{a}}, N_t^{i\tilde{w}a}, N_t^{iw\tilde{a}}, N_t^{iwa}, i). \tag{4.1}$$

The average number of aphids per plant at time t, labeled N_t, equals $\dfrac{\sum_i N_t^i}{M_1 M_2}$. The unit of time t is a day and it takes on integer values between $\{0, t_{max}\}$. In spring, aphids start migrating from buckthorn to soybean [123]. Suppose the initial aphid count migrating to soybean field is lognormally distributed with mean μ and standard error σ i.e.

$$N_0 \sim LogNormal(\mu, \sigma). \tag{4.2}$$

The migrants are all alate female adults [123]. Therefore, at time t=0, the state of any cell is $S(i|t = 0)$= $(0, 0, 0, N_0, i)$. Parameters μ and σ are derived from the mean and variance of N_0 in Appendix C.1. The number of plants infested at the beginning is based on these parameters. The location of infested plants is random. Many species of aphids use visual cues [31] and show edge effects [70] when choosing their host plant. But, evidence on edge effects for soybean aphids is dubious [141, 123]. There are no visual cues for aphids in our model since we assume homogeneity within plants. As a result, we assume their landing on soybean plant is random. Once in the field, soybean aphid may reproduce, grow wings, migrate and die. Each of these biological processes are affected by external factors such as temperature, photo-period and presence of natural enemies which are all assumed to be constant unless stated otherwise.

4.3.1.1 Birth and Death

In section 4.2, we discuss that soybean aphid population is density dependent. This can be either due to bottom-up or top-down factors or a combination of the two. Some of these factors include decline in host plant quality, crowding and presence of natural enemies [126, 32, 25, 41, 26, 101]. Costamagna *et al.* [24] and Matis *et al.* [100] propose two different models in soybean aphid literature which characterize density dependence adequately. The first one is an exponential growth model with decreasing rate of growth to capture bottom-up effects on soybean aphid abundance. The second growth model assumes the death rate to be a function of cumulative aphid density. In our model, we need to assess the nymph and adult aphid density in each grid cell. So, we use the second growth model to represent population dynamics of soybean aphids. The growth

[3]Note that nymphs in soybean aphid literature is used to address all instar stages and therefore a nymphal stage may last 3-4 days [158]. In this paper, we do not track the age of aphids for computational simplicity and so we modify the meaning of the terms nymph and adult aphid.

[4]We suppress a superscript notation in a variable when the notational variable has been summed over.

equation proposed by Matis *et al.* is

$$\frac{dX_t}{dt} = (\lambda - \delta F_t)X_t \tag{4.3}$$

where X_t and $F_t = \int X_t dt$ are the population and cumulative population density respectively at time t. λ and δF_t are the birth and death rate at time t respectively. Therefore, at any time t, total number of new births is λX_t while the number of deaths equal $\delta F_t X_t$. To apply this growth equation in our context, we discretize time i.e. $dt \to 1$. Then soybean aphid population in time $t+1$ for cell i is given by

$$N^i_{t+1} = (1 + \lambda - \delta F^i_t)N^i_t. \tag{4.4}$$

4.3.1.2 Dispersal

Aphids are mostly sedentary [70] because dispersal requires energy which may lower their chance of finding new hosts [70, 145].[5] However, they do move both long and short distance. Their propensity to move and the distance covered depend on their morphs. While aptera walk short distances and usually end up on neighboring plants [7, 70], alates can fly for miles given the right environmental conditions [116]. Aphid dispersal is also a function of their age. Older aphids are less likely to relocate [19]. Long distance movement, also called migration, is quite infrequent compared to the localized displacement [70, 116, 95]. Short distance dispersals or the 'appetitive dispersals' are generally induced by factors such as worsening predator density at the host plants, host plant quality and crowding [64, 126, 70]. In this paper, we only focus on the localized dispersal of aphids because it is the more important of the two types of movement in order to understand their infestation pattern in a field.

To simplify our model, we make the following assumptions about aphid dispersal. (1) Only adult aphids ($l = a$) can move between plants.[6] (2) No mortality occurs during dispersal. This is assumed because of the lack of estimates or data in the literature on death rate during displacement. (3) At any given time period t, soybean aphid can undertake only one interplant trip. (4) The percentage of aphids dispersing/emigrating from a plant at any time t will depend on its aphid density. Donaldson *et al.* [32] regressed the percentage of aphids dispersing from their host plant on the host plant's aphid density and found a positive relationship between the two. But their finding was mostly driven by two outlier plants with aphid density above 4,000.[7] Emigration rates on majority plants did not show an overall increase with aphid density and was 4% for densities less than 4,000 soybean aphid per plant. Let E^{ijh}_t denote the percentage of

[5] Flight not only lowers the longevity of alate aphids but also their fecundity and reproductive period [159].

[6] Remember that in our model, aphids become adults from the day after their birth. Therefore, they are capable of interplant movement when only a day old. In reality, aphids hardly move until fourth instar stage [19], which occurs 4 to 6 days after their birth.

[7] Donaldson *et al.* [32] state the two plants with final aphid density of 6,550 and 10,100 aphids exhibited 20% emigration rate

aphid morph j which emigrate from cell i to h at time t. Then assumption 4 implies that the percentage of total aphid population emigrating from cell i, labeled E_t^i, is a function of N_t^i i.e.

$$\begin{aligned} E_t^i &= \frac{\sum_j N_t^{ij} \sum_h E_t^{ijh}}{N_t^i} \\ &= \varepsilon(N_t^i). \end{aligned} \tag{4.5}$$

Propensity to emigrate is different for the two morphs.[8] Apterous soybean aphids are inactive and generally do not move once they have settled in a desirable feeding location [70] unless instigated by external factors such as deteriorating host plant quality, presence of predators, crowding, wind and rainfall [153, 102, 32, 19, 59, 84]. While apterous adults can only walk, winged adult aphids both walk and fly [19, 94]. Lombaert *et al.* [94] estimated the rate of dispersal for melon aphids to study the impact of their dispersal on their fitness. They found that the rate of walking was 2% and independent of aphid density. But, aerial dispersal was strongly density dependent and it happened in large numbers only when local crowding increased dramatically. Donaldson *et al.* [32], however found that for any given aphid density, the percentage of alates emigrating were at most 4%.

The destination cell h is determined based on the distance aphids can move. Alates can fly miles under right environmental conditions [116]. Zhang *et al.* [158] show that more than 80% of the young adult aphids (12- 72 hours after molting) flew over 0.5km using flight mill experiment.[9] Aptera aphids, on the other hand, avoid movement even under mechanical perturbations [84]. They generally do not walk beyond neighboring plants and therefore are not considered the main force behind enlarging the infection foci beyond the field [69, 149]. We define the neighborhood of aphid type j at cell i, denoted by $\mathcal{N}(j|i)$, as the set of cells which it can migrate to from its host cell. Suppose that aphids always move to a new plant when migrating i.e.

$$h \in \mathcal{N}(j|i) - \{i\}. \tag{4.6}$$

In our model, $E_t^{ijh} > 0$ if and only if $h \in \mathcal{N}(j|i) - \{i\}$ for aphid j. As we have already discussed above, aphids do not have preference about their landing within their neighborhood. In other words, every cell h in the set $\mathcal{N}(j|i) - \{i\}$ is equally likely to be chosen as a destination by an emigrating aphid of type j.

[8] Boiteau [19] studied the relative ability of apterous and alate morphs to disperse from one potato leaflet to another. As per this study, the percentage of buckthorn, potato and green peach wingless adult aphids relocating can be as high as 25%, 45% and 15% respectively depending on their age. On the other hand, these numbers for the winged adults of the respective species were 45%, 40% and 40%.

[9] The older ones ($>$ 72 hours after molting) showed reduced flight time, distance and speed. They also find 16-28°C to be the optimal temperature range for flight.

4.3.1.3 Wing Induction

Various factors such as crowding, host plant quality, presence of natural enemies and photoperiod can induce wings in aphids [87, 110, 135, 73, 125]. Wing induction in aphid can be postnatal or prenatal. Nymphs, in their early instar stages, may grow wings to escape any danger and find a more suitable host plant. On the other hand, if an aphid experience life threatening environmental conditions, she may pass on this information to its successors genetically. As a result, aphids develop wings prenatally. Prenatal wing induction can continue for multiple generations depending on the intensity of environmental factors experienced by their predecessor. This process is difficult to model given our computational limitations and we assume that wing induction is only postnatal.

We also assume that the factor driving postnatal wing induction is crowding.[10] The larger the population density at the host plant, the more likely its nymphs are to develop into alates. Suppose the probability of aphids developing wings, $p(j = w|t, i)$, at t increases at a decreasing rate with population density N_t^i at host plant i [135]. Let this relationship be a logit function of the form

$$\ln\left[\frac{p(j = w|t, i)}{1 - p(j = w|t, i)}\right] = a_0 + a_1 N_t^i. \tag{4.7}$$

Based on the above biological processes, our model can estimate density of nymphs and adults aphids for both morphs $\forall i, j$ at any given time period. At time t, aphid population at cell i is given by

$$N_t^i = \lambda N_{t-1}^i + (1 - E_t^i)(1 - \delta F_{t-1}^i) N_{t-1}^i + \sum_z E_t^{zi} N_t^z \tag{4.8}$$

such that $i \in \mathcal{N}(j|z) - \{z\}\ \forall j$. The state transition rules are explained for each morph and age group in Table C.1.1. We now turn to modeling pest management strategies of farmers.

4.3.2 Pest Management and Profit Maximization

Farmer's pest management strategy consists of two processes- scouting and insecticide treatment. Farmers scout weekly once the aphids arrive to assess aphid density in their field. They then treat the entire field if the infestation can cause economic damage. Let the farmers scouting technique be denoted by k.[11] There are two types of scouting technology.

1. Manual Scouting
2. Aerial Scouting

[10] In our model, soybean plant quality is uniform across all plants. Additionally, effect of natural enemies are not modeled explicitly. Hence, we focus on effect of crowding on wing induction.

[11] In the subsequent sections, we use terminologies 'pest management strategy' and 'scouting technique'

Manual scouting, $k = m$, includes weekly full plant count of aphids on 20-30 randomly sampled plants from the whole field. They will treat their field with insecticide if the observed aphid density is at least ψ. The value of ψ is chosen such that it maximizes the expected profit of the farmer.[12] Aerial scouting ($k = u$), on the other hand, is scouting soybean aphids using UAVs. UAV technology today can detect pest population above a certain threshold, labeled ϕ. Farmers fly drones above their field and take snapshots. These images are color coded to distinguish areas of the field where pest population is above ϕ. Therefore, with drones for surveillance, farmers can find what proportion of their field has soybean aphid density at least equal to ϕ. Given this information, they will choose to treat their fields if the number of soybean plants with at least ϕ soybean aphids are greater than κ. As in manual scouting, the value of κ is also chosen such that it maximizes expected profit of the farmer. Note that ψ and κ are economic thresholds for manual scouting ($ET_m = \psi$) and aerial scouting ($ET_u = \kappa$) respectively. While both of them maximize farmer's expected profit for their respective scouting techniques, the unit of ψ is soybean aphid density per plant while that of κ is the number of plants.

Aphid density estimated based on both types of scouting technique are prone to error. When scouting manually, the error can result for two reasons. First, manual scouting is an arduous process [21] and prone to human error. Second, it does not provide complete coverage of the field. In our paper, we model error in manual scouting due to the second reason only. Marston *et al.* [98] report the performance of UAVs in detecting aphid pressure of 250 aphids per plant. They test it for various models with different combination of wavelengths, trained with SVM. They find that probability of type 2 error is 0.1 on average while that of type 1 error range between 0.2 and 0.4.

Let Y^P be the potential yield of soybean per hectare of field, and P_y be the market price of the crop. Farmers know the arrival time of soybean aphids, t=0, with certainty. They scout and treat with insecticide if necessary in time period *'s'* s.t. $s{=}7t \ \forall \ 7t \in \{1, t_{max}\}$. In other words, farmers scout and can treat once a week. Let $L(N_0, I_s^k)$ represent the percentage yield loss due to soybean aphids, which depends on initial observed aphid infestation (N_0), and how often farmers choose to treat. This decision to treat at time s based on scouting technology k is represented by I_s^k. It equals 1 when treatment is done or 0 otherwise. Therefore, I_s^k is

$$I_s^m = \begin{cases} 0 & N_s < \psi \\ 1 & N_s \geq \psi \end{cases} \tag{4.9}$$

when scouting is done manually and

interchangeably.

[12] Currently, this threshold level is recommended to be 250 soybean aphids per plant.

$$I_s^u = \begin{cases} 0 & \sum_i I(N_s^i > \phi) < \kappa \\ 1 & \sum_i I(N_s^i > \phi) \geq \kappa \end{cases} \tag{4.10}$$

when using UAVs respectively where $I(N_s^i > \phi)$ is the indicator function which equals 1 if $N_s^i > \phi$ or 0 otherwise. I_s^k is a function of ET_k and spatial distribution of aphids in the field $\{N_s^1, N_s^2, ..., N_s^{M_1M_2}\}$ at any time s and therefore can be written as $I_s^k(ET_k, N_s^1, N_s^2, ..., N_s^{M_1M_2})$.[13] When the insecticide is sprayed aphid density of the field becomes $N_t(1-\theta)$ where θ is the efficacy of treatment. Hence, soybean aphid population at the beginning of the next day $s+1$ will equal

$$N_{s+1} = N_s[I_s^k(1-\theta) + (1 - I_s^k)] \quad \forall \quad k,\ s{=}7t. \tag{4.11}$$

Percentage yield loss due to soybean aphids increases at a decreasing rate with aphid density.[14] It converges asymptotically to a yield loss of 40% [122] - 50% [152] when aphid pressure is as high as 80,000 cumulative aphid-days (CAD). Soybean aphid literature mostly use a linear [122, 97] or concave function such as a negative exponential [22] to express percentage yield loss as a function of aphid density. But, soybean aphid also has a damage boundary of 4,000 to 5,000 CAD [119]. As a result, there is negligible yield loss when aphid pressure is in the range of 4,000 to 5,000 CAD [122, 81]. Therefore we use a sigmoidal function to express the relation between soybean aphid density and percentage yield loss. Let the percentage yield loss function be

$$L(N_0, I_s^k) = \frac{\alpha_1}{\alpha_2 + e^{\{\alpha_3 CAD + \alpha_4\}}} \tag{4.12}$$

where CAD is Cumulative aphid-days (CAD). Cumulative aphid-days is a measure of aphid abundance over time [60]. When sampling is done weekly, CAD equals

$$CAD = \sum_s \left[\frac{N_{s-7} + N_s}{2}\right] \times 7. \tag{4.13}$$

In practice, CAD is computed based on the samples taken from the field rather than the true population N_s. But we use N_s $\forall$ s instead so that we can estimate the actual loss.

[13] I_s^m can also be written as a function of $\{N_s^1, N_s^2, ..., N_s^{M_1M_2}\}$ because N_s is an unbiased estimate of mean aphid density of the field.

[14] Zilberman *et al.* [89] demonstrate the importance of correctly specifying the damage abatement processes in the estimation of production functions and input productivity. They show that the use of traditional specifications like Cobb-Douglas overestimate the productivity of damage control inputs and underestimate the productivity of other inputs. Traditional specifications also predict that the spread of resistance will lead to reduction in the use of a damage control agent. In contrast, the specification proposed in their paper captures the real phenomenon, i.e. the use of a damage control agent increases in response to resistance and that it will decrease only when resistance is so widespread that alternative measures are most cost effective.

Soybean plants have also developed tolerance to soybean aphids and therefore can withstand insect feeding without incurring excessive yield losses. Prochaska *et al.* [121] in their field trials found yield loss of 13% for soybean genotype KS4202 at a range of 35,000-50,000 CAD. Ragsdale *et al.* [122], on the other hand, reported a much higher yield loss of 24-36% for the same range of aphid pressure. Moreover, Kucharik *et al.* [85] found that at low aphid pressure (2,000 CAD) soybean plants had higher yield through a compensatory photosynthetic effect than plants which did not experience significant aphid pressure. In this paper, we assume no form of host plant resistance or tolerance to soybean aphids.

Two types of costs are involved in pest management. First is the cost of monitoring or sampling pests to get estimates for the level of infestation. This cost is incurred every time the field is monitored and depends on the scouting method. Let the cost of scouting once equal C^k. Therefore the total cost of scouting in a season equals $C^k \times n(s)$ for a given method k where $n(s)$ is the number of times scouting is done. The other cost is that of treatment which includes expenditure on insecticide and its application. This expense equals the product of cost of a single spray, labeled C_I and number of times farmers decide to treat the field. Therefore for scouting method k, cost of treatment equals $C_I \times \sum_s I_s^k$. Farmers will also incur a fixed cost when purchasing UAVs. But, we assume that farmers have free access to UAVs. This is because we analyze profitability of the two techniques in a single growing season. It may take a farmer multiple seasons to cover the cost of buying a drone. Hence, a profit analysis based on a single growing season will underestimate the profitability of UAV based scouting if the cost of purchase is included.

Let the maximum expected profit of the farmers equal $E(\pi_k^*)$ if they adopt scouting technique k. The objective of a risk neutral farmer is then to choose scouting strategy k such that it maximizes their expected profit $E(\pi^*)$ i.e.

$$E(\pi^*) = max\{E(\pi_m^*), E(\pi_u^*|\phi)\}. \tag{4.14}$$

To solve Equation 4.14, we first need to solve for $E(\pi_k^*)$ $\forall$ k=m,u. This requires identifying the optimal ET for each scouting technique k, given the population dynamics of soybean aphid and the yield loss function. The optimization problem to solve for optimal ET for method k can be stated as follows:

$$\begin{aligned}
E(\pi_k^*) &= \max_{ET_k \geq 0} E(\pi_k) \\
s.t.\ \pi_k &= PY^p(1 - L(N_0, I_s^k)) - C^k n(s) - C_I \sum_s I_s^k \\
Log(N_0) &\sim N(\mu, \sigma) \\
\sigma^2 &= 9.152 * \mu^{1.543} \\
N_t^i &= \lambda N_{t-1}^i + (1 - E_t^i)(1 - \delta F_{t-1}^i) N_{t-1}^i + \sum_z E_t^{zi} N_t^z \quad \forall \quad z
\end{aligned}$$

$$
\begin{aligned}
i &\in \mathcal{N}(j|z) - \{z\} \quad \forall j \\
E_t^i &= \varepsilon(N_t^i) \quad \forall i \\
\ln\left[\frac{p(j=w|t,i)}{1-p(j=w|t,i)}\right] &= a_0 + a_1 N_t^i \quad \forall i \\
I_s^k(ET_k, N_s^1, N_s^2, ..., N_s^{M_1M_2}) &\in \{0,1\} \\
N_{s+1} &= N_s[I_s^k(1-\theta) + (1-I_s^k)] \quad \forall \quad k,s \\
CAD &= \sum_s \left[\frac{N_{s-7}+N_s}{2}\right] \times 7 \\
L(N_0, I_s^k) &= \frac{\alpha_1}{\alpha_2 + e^{\{\alpha_3 CAD + \alpha_4\}}} \\
s &= 7t \quad \forall \quad t>0 \\
t &\in \{0,..,t_{max}\}
\end{aligned}
$$

In the next section, we discuss the parametrization of our benchmark model. The discussion follows in the same order as in section 4.3.

4.4 Model Parametrization

Let there be 10 rows and columns in the field i.e. M_1=M_2=10. Therefore the field has a total of 100 grid cells/plants. The intra-row and inter-row space between plants are assumed to be 3 and 30 inches respectively [63]. The time of arrival of soybean aphids, t=0, corresponds to the V1-V4 growth stage of soybean [64] and t=t_{max} corresponds to the beginning of growth stage R5. This is because to control economic damage from soybean aphid, farmers should undertake measures at most by soybean growth stage R5 [157]; any cure or prevention after R5 has negligible impact on their profit. Therefore, t_{max} on average equals 50 [143].

Birth and Death

Matis *et al.* [100] found that the analytical solution to Equation 4.3 is given by

$$N_t = 4N_{max}e^{-bt_{max}(t-T_{max})}[1+e^{-bt_{max}(t-T_{max})}]^{-2}$$

where N_{max}, T_{max} and b are the predicted size and time of the peak count (in days) and relative rate (per day) respectively. Observed field data can be used to estimate N_{max}, T_{max} and b which in turn gives λ and δ of Equation 4.3.[15] They estimated the parameters of Equation 4.3 based on caged experiments. But under a caged setting, environmental conditions are set artificially and at optimum which inflates population growth. For example, in Matis *et al.* [100], the mean peak aphid density recorded was approximately 34,000 and 55,000 per plant for years 2004 and 2005 respectively. But Bannerman *et al.* [16] report mean peak aphid density in open fields to range between 148-1,726.2 aphids per plant. In this paper, to find optimal pest management strategy

[15] According to Matis *et al.* $\lambda = \dfrac{b(d-1)}{d+1}$ and $\delta = \dfrac{b^2}{2N_{max}}$ where $d = e^{(bT_{max})}$.

of farmers, we need a realistic population growth model. So, we choose parameter values accordingly: Let N_{max}, T_{max} and b be normally distributed with parameters 1000 ± 100, 36.29 ± 0.42 and 0.228 ± 0.23 respectively [82]. This implies that the birth rate λ on average equals 0.23 while the death rate parameter δ is 0.000013. Figure C.1.1 shows a sample population generated using the above parameter values.[16] Let N_0 equal 5 aphids per plant. Using equations in section C.2, $\{\mu, \sigma\}$ is then estimated to be $\{0.78, 1.19\}$.

Dispersal

Donaldson *et al.* [32] estimate percentage of aphids which emigrate as a linear function of soybean aphid density. They find

$$E_t^i = \begin{cases} 1.35 + 0.00184 N_t^i, & N_t^i > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4.15}$$

We assume equation 4.5 equals 4.15.[17] Emigrating population comprises of the alate as well as aptera adults. The probability of an alate emigrating is 4% [32] i.e. $E_t^{iw} = 0.04 \ \forall \ i$. Therefore, total number of alate emigrating from any cell i at time t is $E_t^{iw} N_t^{iw}$. Rest of the emigrants which equal $E_t^i N_t^i - E_t^{iw} N_t^{iw}$ are aptera adult aphids. The neighborhood of alates $\mathcal{N}(\tilde{w})$ has a radius of 0.5km [158]. Thus alate adults can move anywhere in the field.[18] Aptera aphids move very short distance especially when there are suitable hosts nearby [14]. They can move at a speed of 5-20cm per minute or run 15-30cm per minute [116]. Hence, we assume their neighborhood $\mathcal{N}(w)$ has a radius of 10 inches only. This implies that aptera adults can move up to 3 adjacent plants within their row.

Wing Induction

Equation 4.7 is estimated based on the data from Hodgson *et al.* [64] using a logit model with zero intercept. It equals

$$\ln\left[\frac{p(j = w|t)}{1 - p(j = w|t)}\right] = 9.38 \times 10^{-5} N_t + e_t$$

where the error term $e_t \sim N(0, 0.213)$. The constant term a_0 is assumed to be zero because wing induction does not happen when the population density on host plant is 0.

Pest Management and Profit Maximization

Potential yield Y^P and Price P of soybean are 4.04 ton/ha and \$220.46/ton [122]. In our model, as already mentioned before, farmers scout and may treat weekly. Since the window of treatment is 50 days ($\{0, t_{max}\}$), farmers scout (and may treat) 7 times i.e.

[16] Additional tables and figures for this chapter are in and named after appendix C and are duly numbered.

[17] The authors do not report the error structure. We, therefore, begin by assuming this process to be definitive.

[18] The area of our field is 7,290 inch squared which equals $4.7 \times 10^{-6}\ km^2$.

n(s)=7. Every Insecticide treatment costs \$35.82/ha (C_I) [74] and has an effectiveness θ of 99% [157]. Johnson *et al.* [74] also report cost of manual scouting to equal \$9.88/ha. We do not have information on the cost of remote sensing and hence, to begin with, we suppose both types of scouting methods monitor a field at the same cost of \$9.88/ha. This implies that per plant scouting is cheaper when using remote sensing. It is because while UAV scans all the plants, only 20-30 random plants from the field are scouted manually every time sampling is done. Suppose \$9.88/ha cost of monitoring is based on sampling 25 plants per hectare. Then manual scouting costs \$9.88/25 per plant and UAV scouting costs \$9.88/100 if the field has 100 plants in total.[19] The average cost of monitoring C^k per unit area and time is also assumed to be homogeneous of degree one in the size of soybean field. However, it may not be true. Matese *et al.* [99] state that when field size increases from 5 to 50 hectares, the operating costs of UAVs increase by a factor of 2.66. On the other hand, when fields are scouted manually, the cost of labor may increase at a decreasing rate because of diminishing marginal returns.

We assume ϕ to be 250 soybean aphids per plant. This tally threshold is based on the current IPM recommended ET of 250 aphids per plant. The probability of type 1 and 2 error for UAV based scouting in detecting 250 aphids per plant is 0.3 and 0.1 respectively. These values are derived from Marston *et al.* [98] and are based on SVM model with combination of wavelengths 780, 1,010 and 720 nm (Model 3).

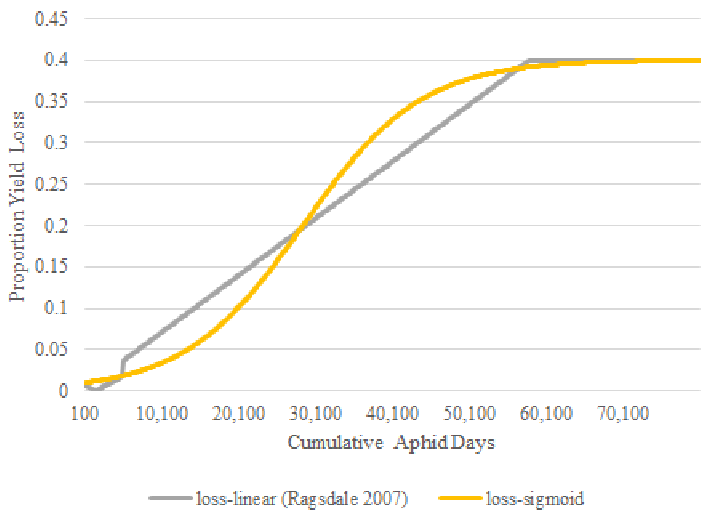


Figure 4.1: Fitted sigmoidal proportional yield loss function versus linear loss function.

Parameters of the yield loss function are chosen such that the damage threshold is $\approx$ 5,000 CAD and yield loss converges to 40% when aphid pressure reaches 80,000 CAD [122, 81]. $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ in Equation 4.12 equals $\{4, 10, 0.00013, 6\}$. Figure 4.1 compares our sigmoidal loss function to the linear loss function based on data from Ragsdale *et al.* [122]. Table C.1.3 puts together the definition of parameters used in our

[19]This implies that when the field is monitored manually using 20 random plants, it costs \$7.9 per hectare.

model, their values and references.[20]

4.5 Simulating the model

The model is initialized at time t=0 by introducing soybean aphid into the field. To do this, we find how many plants are expected to become infested given the distribution of N_0, and their location in the field. Then each of these cells are assigned aphid population based on Equation 4.2. Expected number of infested plants equal probability of infestation times the total number of plants in the field (=100). Probability of infestation is defined as the likelihood that $N_0 \geq 1$. We estimate it using a large sample of N_0's to ensure its unbiasedness. 25,000 samples of N_0 are drawn and fraction of cases where $N_0 \geq 1$ is determined; the probability that a plant will become infested equals this fraction by the law of large numbers. Location of these plants is then determined by randomly drawing $\sum_i I(N_0^i \geq 1)$ grid cells i from a uniform distribution $U(1, 100)$; $I(N_0^i \geq 1)$ is an indicator function that equals 1 if $N_0^i \geq 1 \; \forall \; i$ and 0 otherwise. Next, each of these cells i are infested with N_0^i alate adults through random draws based on Equation 4.2. Each grid cell is also assigned values for n_{max}, T_{max} and b which determine their birth rate λ and death rate parameter δ.

At time $t > 0$, N_t^i is estimated using N_{t-1}^i and Equation 4.4 for every cell i. We then determine the composition of aphid population for each cell i using the transition rules stated in Table C.1.1. Total number of emigrants $E_t^i N_t^i$ is calculated for each cell using Equation 4.15. Out of these emigrants, 4% are alates and rest are aptera adults. These aphids then disperse to cells randomly chosen from within their neighborhood. At each cell, net immigrants from all other cells are added to the existing aphids to get the total number of aphids post dispersal.

At the end of time period t, state vector $S(i|t)$ is known $\forall \; i$. Scouting and treatment is done weekly. On $s = 7t$, 20 cells are randomly chosen from the field to scout manually and average aphid density is estimated. Insecticide treatment follows ($I_s^m = 1$) if observed density is greater than threshold ψ. On the other hand, when remote sensing is used, total number of cells i with $N_t^i > \phi$ is determined. If it is greater than κ, then $I_s^u = 1$. At $t = t_{max}$, cost of treatment ($C_I \sum_s I_s^k$), yield loss proportion $L(N_0, I_s^k)$) and cost of sampling ($C^k n(s)$) are estimated for $k = m, u$. Subsequently, the profit π_k is computed for each scouting strategy k given their respective threshold of spray.

The process explained above presents one full simulation of profit maximization problem based on parameters defined in section 4.4. We execute this model 50 times such that each iteration corresponds to a different draw of initial population N_0. This gives us 50 outcomes (e.g. profits, treatments and losses) for strategy k and its threshold of spray ET_k. The values of ET_u and ET_m are chosen with intervals of 5 and 25 respectively

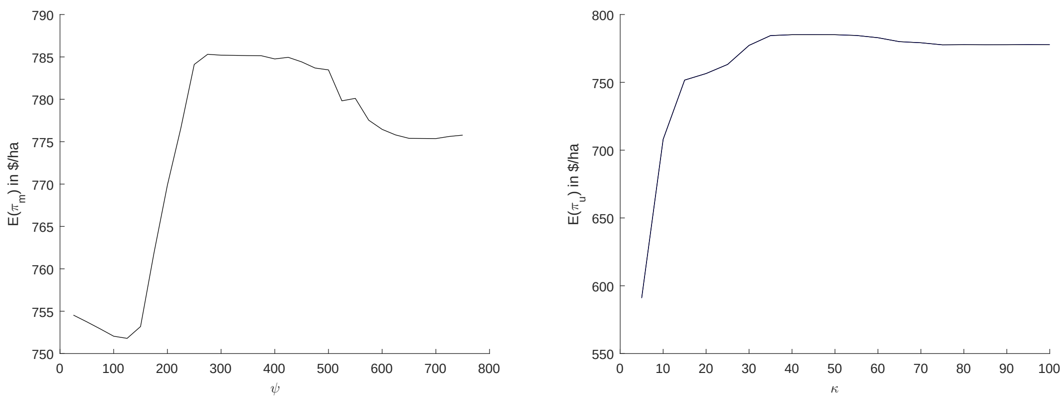
[20] See section C.1.

to keep it computationally light. The expected profit $E(\pi_k)$ is then the average of all 50 profit outcomes. ET_k^* is chosen as the optimal treatment threshold if and only if it gives the maximum expected profit $E(\pi_k^*)$. Once we have ET_m^* and ET_u^*, the optimal pest management strategy and expected profit are inferred using Equation 4.15.

4.6 Results

We first discuss the case of no treatment. The parameter values $\{\mu, \sigma\}$ based on $m = 5$ results in a maximum yield loss of 6.5% if infestation is left untreated, irrespective of the scouting technique. This is equivalent to 15,817 CAD and results in expected profit of \$777.5 per hectare. To put it differently, farmers will end up with at least \$777.5 per hectare if they delay treatment by choosing high κ or ψ in hope of saving the cost of spraying.

Figures 4.2a and 4.2b show how expected profit changes with treatment thresholds for manual scouting (ψ) and UAV based scouting (κ) given that ϕ equals 250 aphids per plant. The tables below each figure gives the probability of insecticide spray at different sampling times, s. The optimal threshold of spray ψ^* equals 275 aphids per plant which gives an expected profit $E(\pi_m^*)$ of \$785.3 per hectare. When scouting manually, farmers choose to treat during $s = 3$ or 4 out of the 7 times. Chances of treatment during these time periods are 6% and 94% respectively. In case of remote sensing, it is optimal to treat the field with insecticide when 45 plants are infested with at least 250 aphids per plant. The optimal treatment threshold κ^* of 45 plants results in treating the field at $s = 4$ and an expected profit of \$785.18 per hectare. Time period $s = 4$ is the 4^{th} week since the beginning of infestation. Because in our model $t = 0$ coincides with V1-V4 stage, $s = 4$ corresponds to R2-R4 reproductive growth stage of soybean [111].



s	1	2	3	4	5	6	7
$E(I_s^m)$	0	0	0.06	0.94	0	0	0

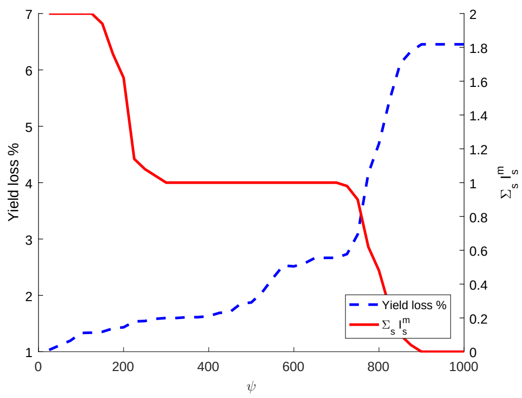
(a) Manual Scouting

s	1	2	3	4	5	6	7
$E(I_s^u)$	0	0	0	1	0	0	0

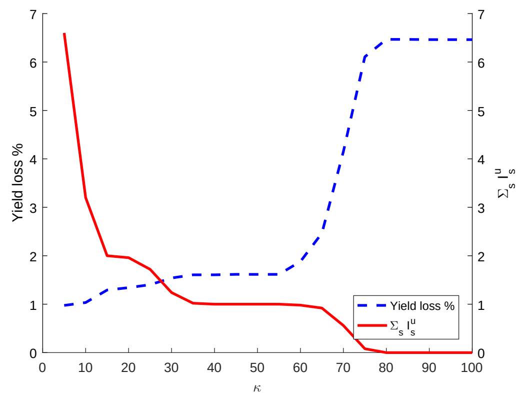
(b) Remote Sensing

Figure 4.2: Expected profit in \$ per hectare using (a) Manual scouting and (b) Remote sensing given ϕ equals 250 aphids per plant.

At their respective optimum, the two techniques have very similar profitability. But figures 4.2a and 4.2b suggest that overall manual scouting is a better option. $E(\pi_m)$ lies between \$750 and \$785 per hectare when soybean aphid pressure is in the range of 0 to 900 aphids per plant. On the contrary, $E(\pi_u|\phi = 250)$ has a minimum of approximately \$600 per hectare when κ equals 5 plants and a maximum of \$785 per hectare at κ=45. Scouting manually estimates aphid pressure more precisely than remote sensing. The economic performance of the two pest management techniques are similar if farmers delay treatment, or in other words, choose high threshold values. Delaying treatment can build up aphid pressure so much that after a point it is no longer useful because of two reasons. First, very high aphid density can lead to irreversible damage in plant quality and yield [142]. Second, density dependence kicks in once aphid pressure reaches n_{max} and causes reduction in infestation level naturally. Figures 4.3a and 4.3b show that at $\psi = 900$ aphids per plant and $\kappa = 100$ plants, percentage yield loss reaches its maximum 6.5% and there are no treatments recommended. As a result, expected profit $E(\pi_m)$ and $E(\pi_u)$ converge to profit under no treatment as shown in figures 4.2a and 4.2b. If they chose to treat early i.e. at low threshold, then manual scouting is significantly more profitable than remote sensing. It is because UAVs are prone to generating false positive outcomes which will cause unnecessary treatments even at low aphid pressure. The probability of type 2 error is 10% and therefore on average 10 out of 100 plants will be falsely reported to have at least 250 aphids. Consequently, at $\kappa = 5$ plants, farmers treat 6.6 times on average out of 7 as shown in Figure 4.3b and reduce their yield loss to 0.97%. But any decision to treat early, say at $\psi = 25$ aphids per plant, based on observations from scouting manually result in 2 treatments in total only while reducing the yield loss to 1.03%.



(a) Manual scouting



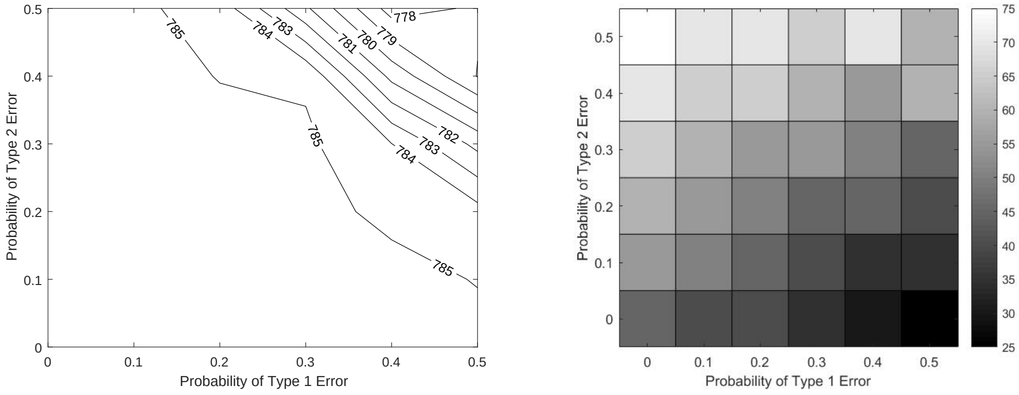
(b) Remote sensing

Figure 4.3: Percentage yield loss and total number of treatments as a function of economic thresholds (a) ψ for manual scouting and (b) κ for remote sensing given $\phi = 250$.

Will improving precision enhance the profitability of UAVs as a scouting technique? Intuitively, expected profits and yield loss should be negatively related with both types of error rates because they comprise the quality of information received from remote sensing. Optimal threshold of treatment should increase when there is higher chance of getting false positives and vice versa. False positive outcomes are equivalent to overestimating aphid pressure. When it becomes more likely, the possibility of unnecessary treatments also increases which in turn will reduce their marginal benefit. As a result, treatment will be delayed in order to increase the marginal benefit of treatment. On the other hand, when type 1 error becomes more probable, treatment threshold should decrease. False negatives imply underestimating the aphid pressure in the field. As a result, the marginal benefit of treatments are high and so they should be conducted at lower threshold. Discerning the effect of probability of type 1 and 2 error on the number of treatments is more complicated. When quality of information goes down, yield loss will increase. This may encourage farmers to treat more times. However, increasing the total number of treatments may not always be needed. Treating early in the season may be enough to suppress the aphid pressure overall and trim yield loss.

To test our hypotheses, we perform sensitivity analysis of pest management decisions with respect to type 1 and 2 error rates. We vary the probability of type 1 and 2 errors from 0 to 0.5 at an interval of 0.05 and estimate the optimal expected profit and treatment threshold for every pair of values.[21] The results are presented in detail in Appendix C.1. Figures 4.4a and 4.4b show how isoprofit lines and treatment threshold vary with the likelihood of type 1 and 2 errors. The isoprofit lines have lower values when the probability of type 1 and type 2 error increased. The relationship between expected profit and these probabilities is negative on average but not in a strict sense. For instance, in Table C.1.4, we see that the expected profit is at a maximum when probabilities of type 1 and type 2 errors are either 0 and 0.5 or 0.1 and 0.4 respectively. Under these conditions, the expected profit is \$785.23 per hectare which is \$0.05 greater than when UAV is 100% accurate. This marginal gain in profit is achieved by reducing yield loss. The gain in profit is also associated with change in treatment decision. Table C.1.5 shows average number of treatments $E(I_s^u)$ $\forall$ s, given the probabilities of type 1 and 2 error. When remote sensing is 100% accurate, profit is maximized by always treating in the third week $s = 3$ at threshold of 45 plants. On the other hand, if UAVs never give false negative information but generate false positives with 0.5 probability, then the optimal strategy is to treat at a threshold of 75 plants 2% and 98% of the times during second and third week respectively. When false positive outcomes are generated 50% of the times, it is impossible to dissociate true case of infestation from the false ones. As a result, treating in 2^{nd} week, even if just 2% of the times, before treating in the 3^{rd} week curbs aphid pressure sufficiently. In this case, spreading the treatment over multiple periods here therefore gives the same results as extra treatments but at a lower cost.

[21] UAVs being systematically right is as informative to farmers as if it is systematically wrong. These conditions are equivalent to when error rates ≤ 0.5 and ≥ 0.5 respectively.



(a) (b)

Figure 4.4: (a) Isoprofit lines and (b) Optimal ET for UAV based scouting as a function of probability of type 1 and 2 errors.

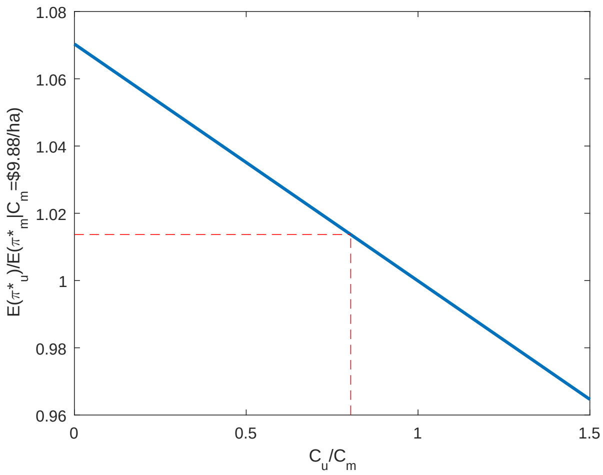


Figure 4.5: Relationship between the ratios of expected optimal profits $E(\pi_u^*)/E(\pi_m^*)$ and monitoring costs C_u/C_m when C_m is fixed at \$9.88 ha^{-1}.

The unmanned aerial systems are expected to become cheaper in near future [118, 156, 33, 35]. Figure 4.5 demonstrates how the cost of monitoring by UAVs will affect their profitability in comparison to manual scouting. We fix the cost of manual scouting C_m at \$9.88 ha^{-1} and express the cost of remote sensing C_u as its fraction. Consequently, $E(\pi_m^*)$ is also fixed at \$785.27 ha^{-1} while $E(\pi_u^*)$ changes with C_u. The expected profit declines linearly with increase in cost of monitoring. It is because the cost of monitoring is independent of economic threshold ET_k. Therefore, any change in it only causes parallel shifts in iso-profit lines and subsequently the optimal expected profit; it does not alter optimal threshold ET_k and treatment decision $I_s^k\ \forall\ k, s$. With our current assumption of $C_u = C_m$, $E(\pi_u^*) \approx E(\pi_m^*)$. If the operating cost of UAVs became negligible (≈ 0), farmers expected profit would rise to \$840.1 ha^{-1} which is 1.07 times more than the optimal profit from manual scouting given its cost C_m. Pederson *et al.* [118] estimated UAV based weed scouting to be 80% as costly as conventional scouting (€19.4 ha^{-1}). If this relationship holds for soybean aphids as well, then $E(\pi_u^*)$ will be \$796.02 ha^{-1}. This amounts to 1.014 times $E(\pi_m^*)$ (refer to the red dotted line in Figure 4.5).

4.7 Conclusion

Soybean farmers have significantly increased their usage of neonicotinoid and foliar insecticides to control soybean aphid infestation. But in the recent years, such treatments are under dispute not only because of the growing resistance in soybean aphids but also due to their harmful effects on pollinators. That is why soybean farmers have been pressured to practice IPM. But many IPM practices, especially manual scouting, are extremely arduous and time consuming, thereby lowering the chances of farmers' adopting them. Since the advancement in remote sensing technology, researchers are exploring UAVs as an alternative for scouting pests including soybean aphids. In this paper, we explore how precise and cheap UAVs have to be to work as a viable scouting option for farmers.

We develop a plant-level bioeconomic model of soybean aphid infestation and control in a field. We analyze alternative scouting techniques which include the traditional manual scouting as well as UAV based surveillance. Our results can be summarized as follow: (i) When farmers treat optimally and on time, their expected profit is approximately \$785 per hectare for both scouting technologies. Yet, UAVs, on average, have greater downside risk than when scouting manually, particularly for farmers who prefer early precautionary treatments. UAVs are higher at risk of detecting false positives which can make farmers perform unnecessary treatments. Conversely, delaying the spray of insecticides may eventually damage the crops irreversibly, irrespective of the scouting method adopted. (ii) Currently, remote sensing technologies can diagnose an infestation level of 250 aphids per plant with almost 80% accuracy. We find that with a tally threshold of 250 aphids per plant, there is not much value added in making UAVs more precise. In fact, perfecting their accuracy at the tally threshold of 250 aphids per plant, can only add 1% to farmers' expected profitability than when they leave the infestation in their fields unchecked. (iii) But, if the cost of operating UAVs was lessened, farmers will readily adopt it as a scouting technique. A 20% reduction in its operating cost increases its expected profitability by 1.4%. Additionally, aerial surveillance of the field helps in identifying the hot spots and consequently treating only those areas. Selective treatments will not only be more cost-effective for farmers but also environmentally beneficial.

This paper is one of the early efforts to understand the economic viability of remote sensing technology in pest management and unfolds a plethora of research possibilities. To adequately model the benefits of aerial surveillance, we need to look at the scope of targeted treatments as discussed in the previous paragraph. Furthermore, the weakness of manual scouting (i.e., relatively small sample sizes) is a strength of UAV scouting and the weakness of UAV scouting (i.e., relatively large measurement error) is a strength of manual scouting. This brings us to another interesting question: Can UAV scouting profitability complement manual scouting rather than substitute for it?

Bibliography

[1] Quick Stats National Agricultural Statistics Service. http://www.nass.usda.gov/. Accessed: Jan, 2016.

[2] United States Department of Agriculture ipmPIPE Restricted Site. https://ipmpipe.zedxinc.com/login.html. Accessed: May, 2018.

[3] Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[4] Charles Abraham and Paschal Sheeran. Acting on intentions: The role of anticipated regret. *British Journal of social psychology*, 42(4):495–511, 2003.

[5] Icek Ajzen. The theory of planned behaviour: reactions and reflections, 2011.

[6] Tavvs M Alves, Ian V Macrae, and Robert L Koch. Soybean aphid (Hemiptera: Aphididae) affects soybean spectral reflectance. *Journal of Economic Entomology*, 108(6):2655–2664, 2015.

[7] Andrei Alyokhin and Gary Sewell. On-soil movement and plant colonization by walking wingless morphs of three aphid species (Homoptera: Aphididae) in greenhouse arenas. *Environmental Entomology*, 32(6):1393–1398, 2003.

[8] Krister P Andersson, Nathan J Cook, Tara Grillos, Maria Claudia Lopez, Carl F Salk, Glenn D Wright, and Esther Mwangi. Experimental evidence on payments for forest commons conservation. *Nature Sustainability*, 1(3):128–135, 2018.

[9] Judd Antin and Aaron Shaw. Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2925–2934, 2012.

[10] Nigel M Asquith, Maria Teresa Vargas, and Sven Wunder. Selling two environmental services: in-kind payments for bird habitat and watershed protection in Los Negros, Bolivia. *Ecological Economics*, 65(4):675–684, 2008.

[11] Shady S Atallah, Miguel I Gómez, Jon M Conrad, and Jan P Nyrop. A plant-level, spatial, bioeconomic model of plant disease diffusion and control: Grapevine leafroll disease. *American Journal of Agricultural Economics*, 97(1):199–218, 2014.

[12] Ofira Ayalon, Tal Goldrath, Gad Rosenthal, and Michal Grossman. Reduction of plastic carrier bag use: An analysis of alternatives in Israel. *Waste Management*, 29(7):2025–2032, 2009.

[13] Christine A Bahlai, Ross M Weiss, and Rebecca H Hallett. A mechanistic model for a tritrophic interaction involving soybean aphid, its host plants, and multiple natural enemies. *Ecological Modelling*, 254:54–70, 2013.

[14] Stephanie M Bailey, Michael E Irwin, Gail E Kampmeier, Catherine E Eastman, and Adrianna D Hewings. Physical and biological perturbations: their effect on the movement of apterous Rhopalosiphum padi (Homoptera: Aphididae) and localized spread of barley yellow dwarf virus. *Environmental Entomology*, 24(1):24–33, 1995.

[15] Sreekala Bajwa, John Rupe, and Johnny Mason. Soybean disease monitoring with leaf reflectance. *Remote Sensing*, 9(2):127, 2017.

[16] JA Bannerman, BP McCornack, DW Ragsdale, N Koper, and AC Costamagna. Predators and alate immigration influence the season-long dynamics of soybean aphid (Hemiptera: Aphididae). *Biological Control*, 117:87–98, 2018.

[17] Anton Bekkerman, Barry K Goodwin, and Nicholas E Piggott. Spatio-temporal risk and severity analysis of soybean rust in the United States. *Journal of Agricultural and Resource Economics*, pages 311–331, 2008.

[18] Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678, 2006.

[19] Gilles Boiteau. Comparative propensity for dispersal of apterous and alate morphs of three potato-colonizing aphid species. *Canadian Journal of Zoology*, 75(9):1396–1403, 1997.

[20] Leah L Bremer, Kathleen A Farley, and David Lopez-Carr. What factors influence participation in payment for ecosystem services programs? an evaluation of Ecuador's sociopáramo program. *Land use policy*, 36:122–133, 2014.

[21] A de F Bueno, Marcelo José Batistela, RCO de F Bueno, Neto França, Marcelo Akita Naime Nishikawa, Adeone Liberio Filho, et al. Effects of integrated pest management, biological control and prophylactic use of insecticides on the management and sustainability of soybean. *Crop Protection*, 30(7):937–945, 2011.

[22] Michael A Catangui, Eric A Beckendorf, and Walter E Riedell. Soybean aphid population dynamics, soybean yield loss, and development of stage-specific economic injury levels. *Agronomy Journal*, 101(5):1080–1092, 2009.

[23] Elijah R Cetas and Maï Yasué. A systematic review of motivational values and conservation success in and around protected areas. *Conservation Biology*, 31(1):203–212, 2017.

[24] AC Costamagna, W Van der Werf, FJJA Bianchi, and DA Landis. An exponential growth model with decreasing r captures bottom-up effects on the population growth of *Aphis glycines* Matsumura (Hemiptera: Aphididae). *Agricultural and Forest Entomology*, 9(4):297–305, 2007.

[25] Alejandro C Costamagna and Douglas A Landis. Predators exert top-down control of soybean aphid across a gradient of agricultural management systems. *Ecological Applications*, 16(4):1619–1628, 2006.

[26] Alejandro C Costamagna, Brian P McCornack, and David W Ragsdale. Within-plant bottom-up effects mediate non-consumptive impacts of top-down control of soybean aphids. *PLoS one*, 8(2):e56394, 2013.

[27] Giovanna d'Adda. Motivation crowding in environmental protection: Evidence from an artefactual field experiment. *Ecological Economics*, 70(11):2083–2097, 2011.

[28] Jeffrey A Davis, Edward B Radcliffe, and David W Ragsdale. Soybean aphid, Aphis glycines Matsumura, a new vector of Potato virus Y in potato. *American Journal of Potato Research*, 82(3):197–201, 2005.

[29] Jan De Leeuw, Patrick Mair, et al. Gifi methods for optimal scaling in r: The package homals. *Journal of Statistical Software*, 31(4):1–20, 2009.

[30] Edward L Deci, Richard Koestner, and Richard M Ryan. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, 125(6):627, 1999.

[31] A. F. G. Dixon. *Host specificity and speciation*, pages 27–58. Springer Netherlands, Dordrecht, 1985.

[32] Jack R Donaldson, Scott W Myers, and Claudio Gratton. Density-dependent responses of soybean aphid (Aphis Glycines Matsumura) populations to generalist predators in mid to late season soybean fields. *Biological Control*, 43(1):111–118, 2007.

[33] Reza Ehsani and Joe Mari Maja. The rise of small UAVs in precision agriculture. *Resource Magazine*, 20(4):18–19, 2013.

[34] D Epstein, JL Frazier, M Purcell-Miramontes, K Hackett, R Rose, T Erickson, T Moriarty, and T Steeger. USDA: Report on the national stakeholders conference on honey bee health. *accessed March 2013*, 2013.

[35] Hannes Eugster and Stephan Nebiker. Geo-registration of video sequences captured from Mini UAVs: Approaches and accuracy assessment. In *Mobile Mapping Technologies Symposium MMT*, volume 1000, pages 1–8, 2007.

[36] LL Evetts and OC Burnside. Effect of early competition on growth of common milkweed. *Weed Science*, 23(1):1–3, 1975.

[37] Driss Ezzine-de Blas, Esteve Corbera, and Renaud Lapeyre. Payments for environmental services and motivation crowding: towards a conceptual framework. *Ecological Economics*, 156:434–443, 2019.

[38] Paul Fackler and Krishna Pacifici. Addressing structural and observational uncertainty in resource management. *Journal of Environmental Management*, 133:27–36, 2014.

[39] Paul L Fackler and Robert G Haight. Monitoring as a partially observable decision problem. *Resource and Energy Economics*, 37:226–241, 2014.

[40] Weimiao Fan and Zheng Yan. Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, 26(2):132–139, 2010.

[41] Tyler B Fox, Douglas A Landis, Fernando F Cardoso, and Christina D Difonzo. Impact of predation on establishment of the soybean aphid, *Aphis glycines* in soybean, *Glycine max*. *Biocontrol*, 50(4):545–563, 2005.

[42] Bruno S Frey and Felix Oberholzer-Gee. The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review*, 87(4):746–755, 1997.

[43] Bruno S Frey and Alois Stutzer. Environmental morale and motivation. *Available at SSRN 900370*, 2006.

[44] R Karina Gallardo and Qianqian Wang. Willingness to pay for pesticides' environmental features and social desirability bias: The case of apple and pear growers. *Journal of Agricultural and Resource Economics*, pages 124–139, 2013.

[45] GAMS. General algebraic modeling system. *GAMS Development Corporation, Release 24.6. Washington, DC, USA*, 2015.

[46] Luis Rico García-Amado, Manuel Ruiz Pérez, and Sara Barrasa García. Motivation for conservation: Assessing integrated conservation and development projects and payments for environmental services in La Sepultura Biosphere Reserve, Chiapas, Mexico. *Ecological Economics*, 89:92–100, 2013.

[47] Uri Gneezy et al. The w effect of incentives. *University of Chicago Graduate School of Business*, 2003.

[48] Uri Gneezy and Aldo Rustichini. Pay enough or don't pay at all. *The Quarterly journal of Economics*, 115(3):791–810, 2000.

[49] HCJ Godfray, Tjeerd Blacquiere, Linda M Field, Rosemary S Hails, Simon G Potts, Nigel E Raine, Adam J Vanbergen, and Angela R McLean. A restatement of recent advances in the natural science evidence base concerning neonicotinoid insecticides and insect pollinators. *Proceedings of the Royal Society B: Biological Sciences*, 282(1818):20151821, 2015.

[50] Dave Goulson. An overview of the environmental risks posed by neonicotinoid insecticides. *Journal of Applied Ecology*, 50(4):977–987, 2013.

[51] Pamela Grimm. Social desirability bias. *Wiley International Encyclopedia of Marketing*, 2010.

[52] Robert G Haight and Stephen Polasky. Optimal control of an invasive species with imperfect information about the level of infestation. *Resource and Energy Economics*, 32(4):519–533, 2010.

[53] Rebecca H Hallett, Christine A Bahlai, Yingen Xue, and Arthur W Schaafsma. Incorporating natural enemy units into a dynamic action threshold for the soybean aphid, Aphis glycines (Homoptera: Aphididae). *Pest Management Science*, 70(6):879–888, 2014.

[54] Anthony A Hanson, James Menger-Anderson, Celia Silverstein, Bruce D Potter, Ian V MacRae, Erin W Hodgson, and Robert L Koch. Evidence for soybean aphid (Hemiptera: Aphididae) resistance to pyrethroid insecticides in the upper midwestern United States. *Journal of Economic Entomology*, 110(5):2235–2246, 2017.

[55] Nicholle M Hatton. *Use of small unmanned aerial system for validation of sudden death syndrome in soybean through multispectral and thermal remote sensing.* PhD thesis, Kansas State University, 2018.

[56] Mickaël Henry, Maxime Beguin, Fabrice Requier, Orianne Rollin, Jean-François Odoux, Pierrick Aupinel, Jean Aptel, Sylvie Tchamitchian, and Axel Decourtye. A common pesticide decreases foraging success and survival in honey bees. *Science*, 336(6079):348–350, 2012.

[57] Ittai Herrmann, Steven Vosberg, Prabu Ravindran, Aditya Singh, Hao-Xun Chang, Martin Chilvers, Shawn Conley, and Philip Townsend. Leaf and canopy level detection of Fusarium virguliforme (sudden death syndrome) in soybean. *Remote Sensing*, 10(3):426, 2018.

[58] Don Hershman. Personal communication. University of Kentucky Research and Education Center. 2013-02-03.

[59] Chris Hodgson. Dispersal of apterous aphids (Homoptera: Aphididae) from their host plant and its significance. *Bulletin of Entomological Research*, 81(4):417–427, 1991.

[60] Erin Hodgson, Greg VanNostrand, and Josh Sievers. Soybean aphid efficacy evaluation in northwest Iowa. *Farm Progress Reports*, 2015(1):138, 2016.

[61] Erin W Hodgson, BP McCornack, KA Koch, DW Ragsdale, KD Johnson, ME O'Neal, EM Cullen, HJ Kraiss, CD DiFonzo, and LM Behnken. Field validation of speed scouting for soybean aphid. *Crop Management*, 6(1):0–0, 2007.

[62] Erin W Hodgson, BP McCornack, K Tilmon, and JJ Knodel. Management recommendations for soybean aphid (Hemiptera: Aphididae) in the United States. *Journal of Integrated Pest Management*, 3(1):E1–E10, 2012.

[63] EW Hodgson, EC Burkness, WD Hutchison, and DW Ragsdale. Enumerative and binomial sequential sampling plans for soybean aphid (Homoptera: Aphididae) in soybean. *Journal of Economic Entomology*, 97(6):2127–2136, 2004.

[64] EW Hodgson, RC Venette, M Abrahamson, and DW Ragsdale. Alate production of soybean aphid (Homoptera: Aphididae) in Minnesota. *Environmental Entomology*, 34(6):1456–1463, 2005.

[65] Elizabeth E Holmes, Mark A Lewis, JE Banks, and RR Veit. Partial differential equations in ecology: Spatial interactions and population dynamics. *Ecology*, 75(1):17–29, 1994.

[66] Darrell Hueth and Uri Regev. Optimal agricultural pest management with increasing pest resistance. *American Journal of Agricultural Economics*, 56(3):543–552, 1974.

[67] Terrance Hurley and Paul Mitchell. Value of neonicotinoid seed treatments to US soybean farmers. *Pest Management Science*, 73(1):102–112, 2017.

[68] ipmPIPE. ipm Pest Information Platform for Extension and Education. http://www.ipmpipe.org/. Accessed: Jan, 2016.

[69] ME Irwin and JM Thresh. Epidemiology of barley yellow dwarf: a study in ecological complexity. *Annual Review of Phytopathology*, 28(1):393–424, 1990.

[70] Michael E Irwin, Gail E Kampmeier, and Wolfgang W Weisser. Aphid movement: process and consequences. *Aphids as Crop Pests*, pages 153–186, 2007.

[71] Scott A Isard, Stuart H Gage, Paul Comtois, and Joseph M Russo. Principles of the atmospheric pathway for invasive species applied to soybean rust. *AIBS Bulletin*, 55(10):851–861, 2005.

[72] Robert C Johansson, Michael J Livingston, John Westra, and Kurt Guidry. Simulating the US impacts of alternative Asian soybean rust treatment regimes. *Agricultural and Resource Economics Review*, 35(1):116–127, 2006.

[73] Bruce Johnson. Wing polymorphism in aphids ii. interaction between aphids. *Entomologia Experimentalis et Applicata*, 8(1):49–64, 1965.

[74] Kevin D Johnson, Matthew E O'Neal, David W Ragsdale, Christina D Difonzo, Scott M Swinton, Philip M Dixon, Bruce D Potter, Erin W Hodgson, and Alejandro C Costamagna. Probability of cost-effective management of soybean aphid (Hemiptera: Aphididae) in North America. *Journal of Economic Entomology*, 102(6):2101–2108, 2009.

[75] David Kaczan, Brent M Swallow, et al. Designing a payments for ecosystem services (PES) program to reduce deforestation in Tanzania: An assessment of payment approaches. *Ecological Economics*, 95:20–30, 2013.

[76] AJ Karley, WE Parker, JW Pitchford, and AE Douglas. The mid-season crash in aphid populations: why and how does it occur? *Ecological Entomology*, 29(4):383–388, 2004.

[77] Fraser Kent. Glyphosate Resistant Weeds – Intensifying. *Stratus Ag Research*, 2013. Accessed:June, 2020.

[78] John Kerr, Mamta Vardhan, and Rohit Jindal. Prosocial behavior and incentives: evidence from field experiments in rural Mexico and Tanzania. *Ecological Economics*, 73:220–227, 2012.

[79] Robert Koch. Soybean Growth Stages. https://extension.umn.edu/soybean-pest-management/scouting-soybean-aphid. Accessed:2019.

[80] Robert L Koch, Erin W Hodgson, Janet J Knodel, Adam J Varenhorst, and Bruce D Potter. Management of insecticide-resistant soybean aphids in the Upper Midwest of the United States. *Journal of Integrated Pest Management*, 9(1):23, 2018.

[81] Robert L Koch, Bruce D Potter, Phillip A Glogoza, Erin W Hodgson, Christian H Krupke, John F Tooker, Christina D DiFonzo, Andrew P Michel, Kelley J Tilmon, Travis J Prochaska, et al. Biology and economics of recommendations for insecticide-based management of soybean aphid. *Plant Health Progress*, 17(4):265–269, 2016.

[82] Koch, Robert L., University of Minnesota. Personal communication.

[83] Nicolas Kosoy, Miguel Martinez-Tuna, Roldan Muradian, and Joan Martinez-Alier. Payments for environmental services in watersheds: Insights from a comparative study of three cases in Central America. *Ecological Economics*, 61(2-3):446–455, 2007.

[84] Caitlin Colleen Krueger. The effects of rainstorms on soybean aphids, *Aphis glycines* (Matsumura). Master's thesis, University of Minnesota, 2014.

[85] Christopher J Kucharik, Amelia C Mork, Timothy D Meehan, Shawn P Serbin, Aditya Singh, Philip A Townsend, Kaitlin Stack Whitney, and Claudio Gratton. Evidence for compensatory photosynthetic and yield response of soybeans to aphid herbivory. *Journal of Economic Entomology*, 109(3):1177–1187, 2016.

[86] Fred Kuchler, Michael Duffy, RD Shrum, WM Dowler, et al. Potential economic consequences of the entry of an exotic fungal pest: The case of soybean rust. *Phytopathology*, 74(8):916–920, 1984.

[87] Li-Hua Lãœ and Rui-Lu Chen. Analysis of factors inducing alatae in Aphis glycines. *Kun chong xue bao. Acta entomologica Sinica*, 36:143–149, 1993.

[88] Yan Li, Curtis B Hill, and Glen L Hartman. Effect of three resistant soybean genotypes on the fecundity, mortality, and maturation of soybean aphid (Homoptera: Aphididae). *Journal of Economic Entomology*, 97(3):1106–1111, 2004.

[89] Erik Lichtenberg and David Zilberman. The econometrics of damage control: why specification matters. *American Journal of Agricultural Economics*, 68(2):261–273, 1986.

[90] Mariëlle Linting, Jacqueline J Meulman, Patrick JF Groenen, and Anita J van der Koojj. Nonlinear principal components analysis: Introduction and application. *Psychological methods*, 12(3):336, 2007.

[91] Michael J Livingston. US soybean producer perceptions and management of soybean rust in the United States under the USDA pest information platform for extension and education. *Agricultural and Resource Economics Review*, 39(3):547–560, 2010.

[92] Mike Livingston, Rob Johansson, Stan Daberkow, Michael Roberts, Mark Ash, and Vince Breneman. Economic and policy implications of wind-borne entry of Asian soybean rust into the United States. *USDA Outlook Report No.(OCS04D02), April, www. ers. usda. gov/publications/OCS/APR04/OCS04D02*, 2004.

[93] Lognormal Distribution. https://www.mathworks.com/help/stats/lognormal-distribution.html. Accessed:2019.

[94] Eric Lombaert, Roger Boll, and Laurent Lapchin. Dispersal strategies of phytophagous insects at a local scale: adaptive potential of aphids in an agricultural environment. *BMC Evolutionary Biology*, 6(1):75, 2006.

[95] Hugh D Loxdale, Jim Hardie, Susan Halbert, Robert Foottit, Neil AC Kidd, and Clive I Carter. The relative importance of short-and long-range movement of flying aphids. *Biological Reviews*, 68(2):291–311, 1993.

[96] Patrick Mair. *Modern psychometrics with R*. Springer, 2018.

[97] Lia Marchi-Werle, Edson LL Baldin, Hillary D. Fischer, Tiffany Heng-Moss, and Thomas Hunt. Economic Injury Levels for Aphis glycines Matsumura (Hemiptera:Aphididae) on the Soybean Aphid Tolerant KS4202 Soybean. *Journal of Economic Entomology*, 110(5):2100–2108, 2017.

[98] Zachary P.D. Marston. *Advancing Remote Sensing For Soybean Aphid (Hemiptera: Aphididae) Management In Soybean*. PhD thesis, University of Minnesota, 2019.

[99] Alessandro Matese, Piero Toscano, Salvatore Di Gennaro, Lorenzo Genesio,

Francesco Vaccari, Jacopo Primicerio, Claudio Belli, Alessandro Zaldei, Roberto Bianconi, and Beniamino Gioli. Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote Sensing*, 7(3):2971–2990, 2015.

[100] James H Matis, Thomas R Kiffe, Wopke van der Werf, Alejandro C Costamagna, Timothy I Matis, and William E Grant. Population dynamics models based on cumulative density dependent feedback: A link to the logistic growth curve and a test for symmetry using aphid data. *Ecological Modelling*, 220(15):1745–1751, 2009.

[101] BP McCornack, DW Ragsdale, and RC Venette. Demography of soybean aphid (Homoptera: Aphididae) at summer temperatures. *Journal of Economic Entomology*, 97(3):854–861, 2004.

[102] Brian P McCornack, AC Costamagna, and DW Ragsdale. Within-plant distribution of soybean aphid (Hemiptera: Aphididae) and development of node-based sample units for estimating whole-plant densities in soybean. *Journal of Economic Entomology*, 101(4):1488–1500, 2008.

[103] Jacqueline J Meulman, Anita J Van der Kooij, and Willem J Heiser. Principal components analysis with nonlinear optimal scaling transformations for ordinal and nominal data. *The Sage Handbook of Quantitative Methodology for the Social Sciences*, pages 49–72, 2004.

[104] George Michailidis and Jan De Leeuw. The gifi system of descriptive multivariate analysis. *Statistical Science*, pages 307–336, 1998.

[105] Jeffrey C Milder, Sara J Scherr, and Carina Bracer. Trends and future potential of payment for ecosystem services to alleviate rural poverty in developing countries. *Ecology and Society*, 15(2), 2010.

[106] HJ Miller. Geocomputation. *The SAGE Handbook of Spatial Analysis. Sage, London*, pages 397–418, 2009.

[107] Sally A Miller, Fen D Beed, and Carrie Lapaire Harmon. Plant disease diagnostic capabilities and networks. *Annual Review of Phytopathology*, 47:15–38, 2009.

[108] Yuichi Mori, Masahiro Kuroda, and Naomichi Makino. *Nonlinear principal component analysis and its applications*. Springer, 2016.

[109] Pallab Mozumder and Robert P Berrens. Social context, financial stakes and hypothetical bias: an induced value referendum experiment. *Applied Economics*, 43(29):4487–4499, 2011.

[110] Christine B Müller, Iain S Williams, and Jim Hardie. The role of nutrition, crowding and interspecific interactions in the development of winged aphids. *Ecological Entomology*, 26(3):330–340, 2001.

[111] Seth L. Naeve. Soybean Growth Stages. https://extension.umn.edu/growing-soybean/soybean-growth-stages#herbicides%3A-How-growth-stage-affects-applications-539863. Accessed:2019.

[112] Ulf Narloch, Unai Pascual, and Adam G Drucker. Collective action dynamics under external rewards: experimental insights from andean farming communities. *World Development*, 40(10):2096–2107, 2012.

[113] Karen Oberhauser, Ruscena Wiederholt, Jay E Diffendorfer, Darius Semmens, Leslie Ries, Wayne E Thogmartin, Laura Lopez-Hoffman, and Brice Semmens. A trans-national monarch butterfly population model and implications for regional conservation priorities. *Ecological Entomology*, 42(1):51–60, 2017.

[114] Matt O'Neal, Kevin Johnson, Erin Hodgson, Dave Ragsdale, Ian MacRae, Bruce Potter, Chris DiFonzo, Kelley Tilmon, Eileen Cullen, Phillip Glogoza, et al. Comment on "soybean aphid population dynamics, soybean yield loss, and development of stage-specific economic injury levels" by ma catangui, ea beckendorf, and we riedell, agron. j. 101: 1080–1092 (2009). *Agronomy Journal*, 102(1):55–59, 2010.

[115] K Ostlie. Soybean aphid reduces yields: harvest results from insecticide strip trials. *University of Minnesota, St. Paul, MN. Online at: http://www.soybeans.umn.edu/crop/insects/aphid/studyresults.htm*, 2001.

[116] Hazel R Parry. Cereal aphid movement: general principles and simulation modelling. *Movement Ecology*, 1(1):14, 2013.

[117] Jacob R Pecenka and Jonathan G Lundgren. Non-target effects of clothianidin on monarch butterflies. *The Science of Nature*, 102(3-4):19, 2015.

[118] Søren Marcus Pedersen, Spyros Fountas, Henrik Have, and BS Blackmore. Agricultural robots—system analysis and economic feasibility. *Precision Agriculture*, 7(4):295–308, 2006.

[119] Larry P Pedigo, Scott H Hutchins, and Leon G Higley. Economic injury levels in theory and practice. *Annual Review of Entomology*, 31(1):341–368, 1986.

[120] John Pleasants. Milkweed restoration in the midwest for monarch butterfly recovery: estimates of milkweeds lost, milkweeds remaining and milkweeds that must be added to increase the monarch population. *Insect Conservation and Diversity*, 10(1):42–53, 2017.

[121] TJ Prochaska, LM Pierson, Edson Luiz Lopes Baldin, TE Hunt, TM Heng-Moss, and John C Reese. Evaluation of late vegetative and reproductive stage soybeans for resistance to soybean aphid (Hemiptera: Aphididae). *Journal of Economic Entomology*, 106(2):1036–1044, 2013.

[122] David W Ragsdale, BP McCornack, RC Venette, BD Potter, Ian V MacRae, Erin W Hodgson, Matthew E O'Neal, Kevin D Johnson, RJ O'neil, CD DiFonzo,

et al. Economic threshold for soybean aphid (Hemiptera: Aphididae). *Journal of Economic Entomology*, 100(4):1258–1267, 2007.

[123] David W Ragsdale, David J Voegtlin, and Robert J O'neil. Soybean aphid biology in North America. *Annals of the Entomological Society of America*, 97(2):204–208, 2004.

[124] Karl M Rich, Alex Winter-Nelson, and Nicholas Brozović. Modeling regional externalities with heterogeneous incentives and fixed boundaries: Applications to foot and mouth disease control in South America. *Review of Agricultural Economics*, 27(3):456–464, 2005.

[125] AF Ríos Martínez and AC Costamagna. Effects of crowding and host plant quality on morph determination in the soybean aphid, Aphis glycines. *Entomologia Experimentalis et Applicata*, 166(1):53–62, 2018.

[126] Aldo F Rios Martinez and Alejandro C Costamagna. Dispersal to predator-free space counterweighs fecundity costs in alate aphid morphs. *Ecological Entomology*, 42(5):645–656, 2017.

[127] Michael J Roberts, David Schimmelpfennig, Michael J Livingston, and Elizabeth Ashley. Estimating the value of an early-warning system. *Review of Agricultural Economics*, 31(2):303–329, 2009.

[128] Michael J Roberts, David E Schimmelpfennig, Elizabeth Ashley, Michael J Livingston, Mark S Ash, Utpal Vasavada, et al. The value of plant disease early-warning systems: a case study of USDA's soybean rust coordinated framework. Technical report, United States Department of Agriculture, Economic Research Service, 2006.

[129] Julian Rode, Erik Gómez-Baggethun, and Torsten Krause. Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence. *Ecological Economics*, 117:270–282, 2015.

[130] Carlos Rodriguez-Sickert, Ricardo Andrés Guzmán, and Juan Camilo Cárdenas. Institutions influence preferences: Evidence from a common pool resource experiment. *Journal of Economic Behavior & Organization*, 67(1):215–227, 2008.

[131] AY Rossman, K Britton, D Luster, M Palm, MH Royer, and J Sherald. Main content area evaluating the threat posed by fungi on the APHIS list of regulated plant pests. *Plant Health Progress*, pages 603–610, 2006.

[132] Shannon Sahlqvist, Yena Song, Fiona Bull, Emma Adams, John Preston, David Ogilvie, iConnect Consortium, et al. Effect of questionnaire length, personalisation and reminder type on response rate to a complex postal survey: randomised controlled trial. *BMC Medical Research Methodology*, 11(1):62, 2011.

[133] James N Sanchirico and James E Wilen. Optimal spatial management of renewable resources: matching policy scope to ecosystem scale. *Journal of Environmental Economics and Management*, 50(1):23–46, 2005.

[134] Brice X Semmens, Darius J Semmens, Wayne E Thogmartin, Ruscena Wiederholt, Laura López-Hoffman, Jay E Diffendorfer, John M Pleasants, Karen S Oberhauser, and Orley R Taylor. Quasi-extinction risk and population targets for the eastern, migratory population of monarch butterflies (danaus plexippus). *Scientific Reports*, 6:23265, 2016.

[135] MJP Shaw. Effects of population density on alienicolae of Aphis fabae Scop. I. The effect of crowding on the production of alatae in the laboratory. *Annals of Applied Biology*, 65(2):191–196, 1970.

[136] Martin D Smith, James N Sanchirico, and James E Wilen. The economics of spatial-dynamic processes: applications to renewable resources. *Journal of Environmental Economics and Management*, 57(1):104–121, 2009.

[137] Steven M Smith. Economic incentives and conservation: Crowding-in social norms in a groundwater commons. *Journal of Environmental Economics and Management*, 90:147–174, 2018.

[138] Michelle J Solensky. Overview of monarch overwintering biology. *The Monarch Butterfly: Biology and Conservation. Cornell University Press, Ithaca, NY*, pages 117–120, 2004.

[139] Matthew Sommerville, EJ Milner-Gulland, Michael Rahajaharison, and Julia PG Jones. Impact of a community-based payment for environmental services intervention on forest use in Menabe, Madagascar. *Conservation Biology*, 24(6):1488–1498, 2010.

[140] Dariane Souza, Bruno C Vieira, Bradley K Fritz, Wesley C Hoffmann, Julie A Peterson, Greg R Kruger, and Lance J Meinke. Western corn rootworm pyrethroid resistance confirmed by aerial application simulations of commercial insecticides. *Scientific reports*, 9(1):1–10, 2019.

[141] Soybean Aphid Management. https://www.canr.msu.edu/field_crops/uploads/archive/Soybean%20aphid%20mgmt.pdf. Accessed:2019.

[142] Soybean Aphid Management. https://www.pioneer.com/us/agronomy/soybean-aphid-mgmt.html. Accessed:2019.

[143] Soybean Growth Stages. https://extension.umn.edu/growing-soybean/soybean-growth-stages#days-between-stages-539862. Accessed:2019.

[144] J Stack, K Cardwell, R Hammerschmidt, J Byrne, R Loria, K Snover-Clift, W Baldwin, G Wisler, H Beck, R Bostock, et al. The national plant diagnostic network. *Plant Disease*, 90(2):128–136, 2006.

[145] Seyed Mohammad Tabadkani, Seyed Mohammad Ahsaei, Vahid Hosseininaveh, and Jamasb Nozari. Food stress prompts dispersal behavior in apterous pea aphids: Do activated aphids incur energy loss? *Physiology & Behavior*, 110:221–225, 2013.

[146] Leigh Tesfatsion and Kenneth L Judd. *Handbook of Computational Economics: Agent-based Computational Economics*, volume 2. Elsevier, 2006.

[147] Wayne E Thogmartin, Laura López-Hoffman, Jason Rohweder, Jay Diffendorfer, Ryan Drum, Darius Semmens, Scott Black, Iris Caldwell, Donita Cotter, Pauline Drobney, et al. Restoring monarch butterfly habitat in the Midwestern US:'all hands on deck'. *Environmental Research Letters*, 12(7):074005, 2017.

[148] H Van den Berg, D Ankasah, A Muhammad, R Rusli, HA Widayanto, HB Wirasto, and I Yully. Evaluating the role of predation in population fluctuations of the soybean aphid Aphis glycines in farmer's fields in Indonesia. *Journal of Applied Ecology*, pages 971–984, 1997.

[149] Helmut F Van Emden and Richard Harrington. *Aphids as crop pests*. Cabi, 2017.

[150] Gert Van Hecken and Johan Bastiaensen. Payments for ecosystem services in Nicaragua: do market-based approaches work? *Development and Change*, 41(3):421–444, 2010.

[151] Grace B Villamor and Meine van Noordwijk. Social role-play games vs individual perceptions of conservation and PES agreements for maintaining rubber agroforests in Jambi (Sumatra), Indonesia. *Ecology and Society*, 16(3), 2011.

[152] Xibei Wang, Yihao Fang, Zhizhong Lin, Lirong Zhang, and Huadi Wang. A study on the damage and economic threshold of the soybean aphid at the seedling stage. *Plant Protection (Institute of Plant Protection, CAAS, China)*, 20:12–13, 1994.

[153] Rebecca Whalen and Jason P Harmon. Rag1 aphid resistant soybeans alter the movement and distribution of soybean aphid (Hemiptera: Aphididae). *Environmental Entomology*, 41(6):1426–1434, 2012.

[154] Byron K Williams. Resolving structural uncertainty in natural resources management using POMDP approaches. *Ecological Modelling*, 222(5):1092–1102, 2011.

[155] Allen Wrather and Steve Koenning. Effects of diseases on soybean yields in the United States 1996 to 2007. *Plant Health Progress.Online*, 2009.

[156] Chunhua Zhang and John M Kovacs. The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, 13(6):693–712, 2012.

[157] Wei Zhang and Scott M Swinton. Incorporating natural enemies in an economic threshold for dynamically optimal pest management. *Ecological Modelling*, 220(9):1315–1324, 2009.

[158] Ying Zhang, Limin Wang, Kongming Wu, Kris AG Wyckhuys, and George E Heimpel. Flight performance of the soybean aphid, Aphis glycines (Hemiptera: Aphididae) under different temperature and humidity regimens. *Environmental Entomology*, 37(2):301–306, 2008.

[159] Ying Zhang, Kongming Wu, Kris AG Wyckhuys, and George E Heimpel. Trade-offs between flight and fecundity in the soybean aphid (Hemiptera: Aphididae). *Journal of Economic Entomology*, 102(1):133–138, 2009.

Appendices

A.1

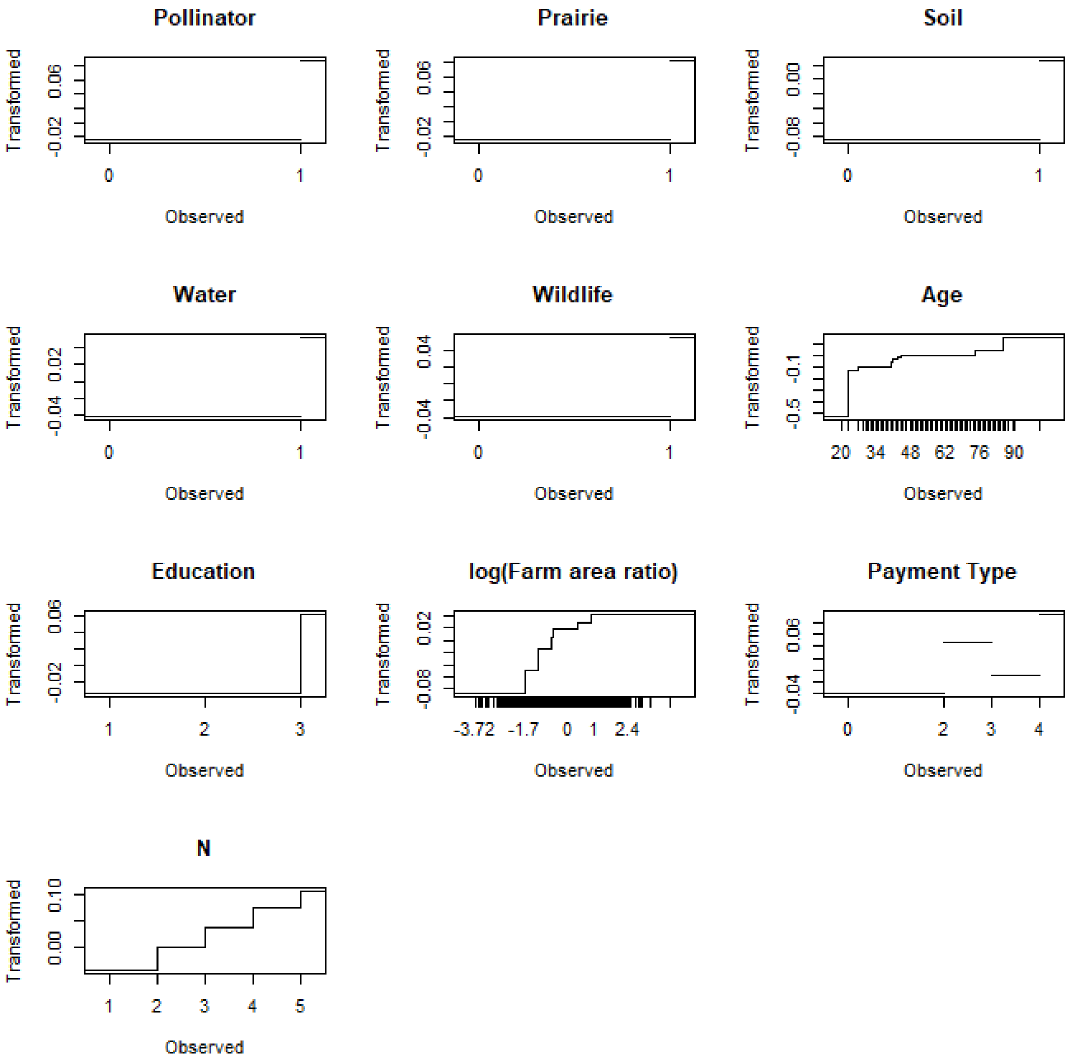


Figure A.1.6: Transformation plots for CATPCA in Method II.

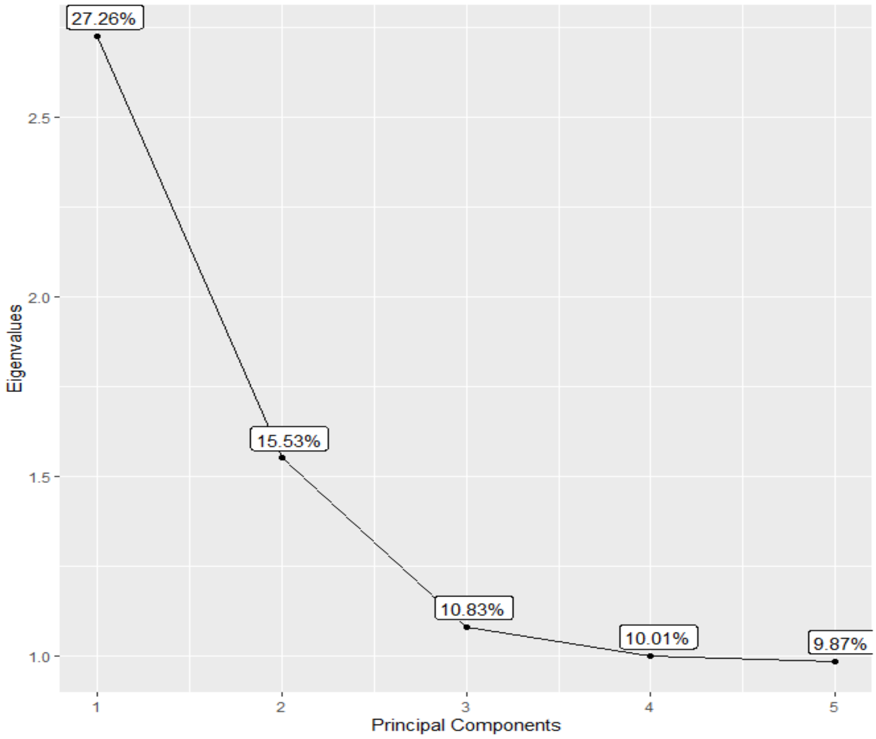


Figure A.1.7: Scree plots for CATPCA in Method II.

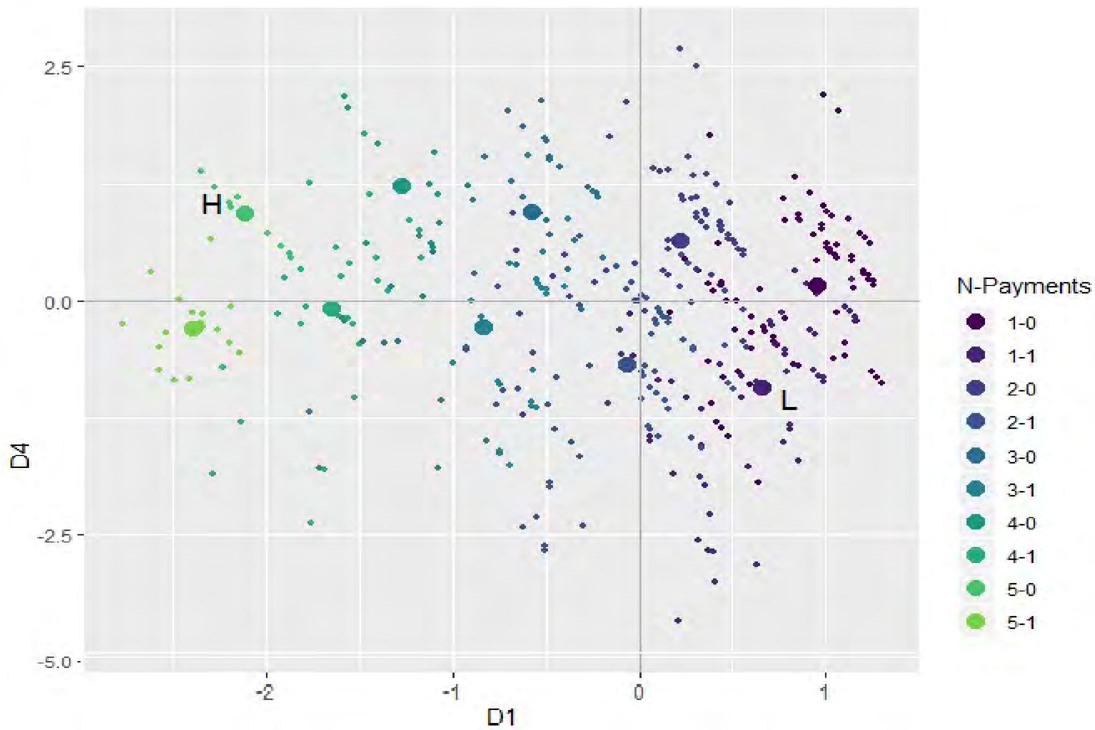


Figure A.1.8: Component scores color coded based on total number of conservation activities and type of conservation payment. The axes D1 and D4 are the principal components from CATPCA in Method II.

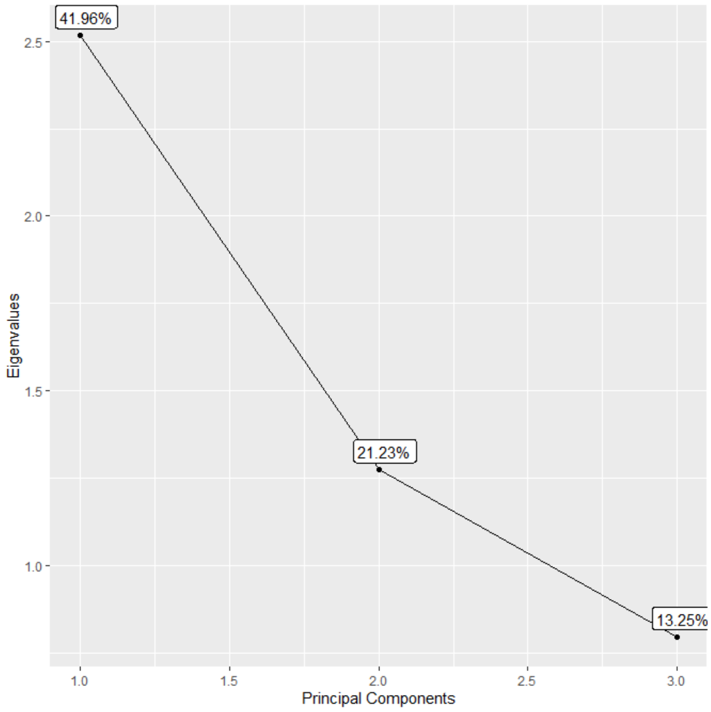


Figure A.1.9: Scree plots for CATPCA in Method III.

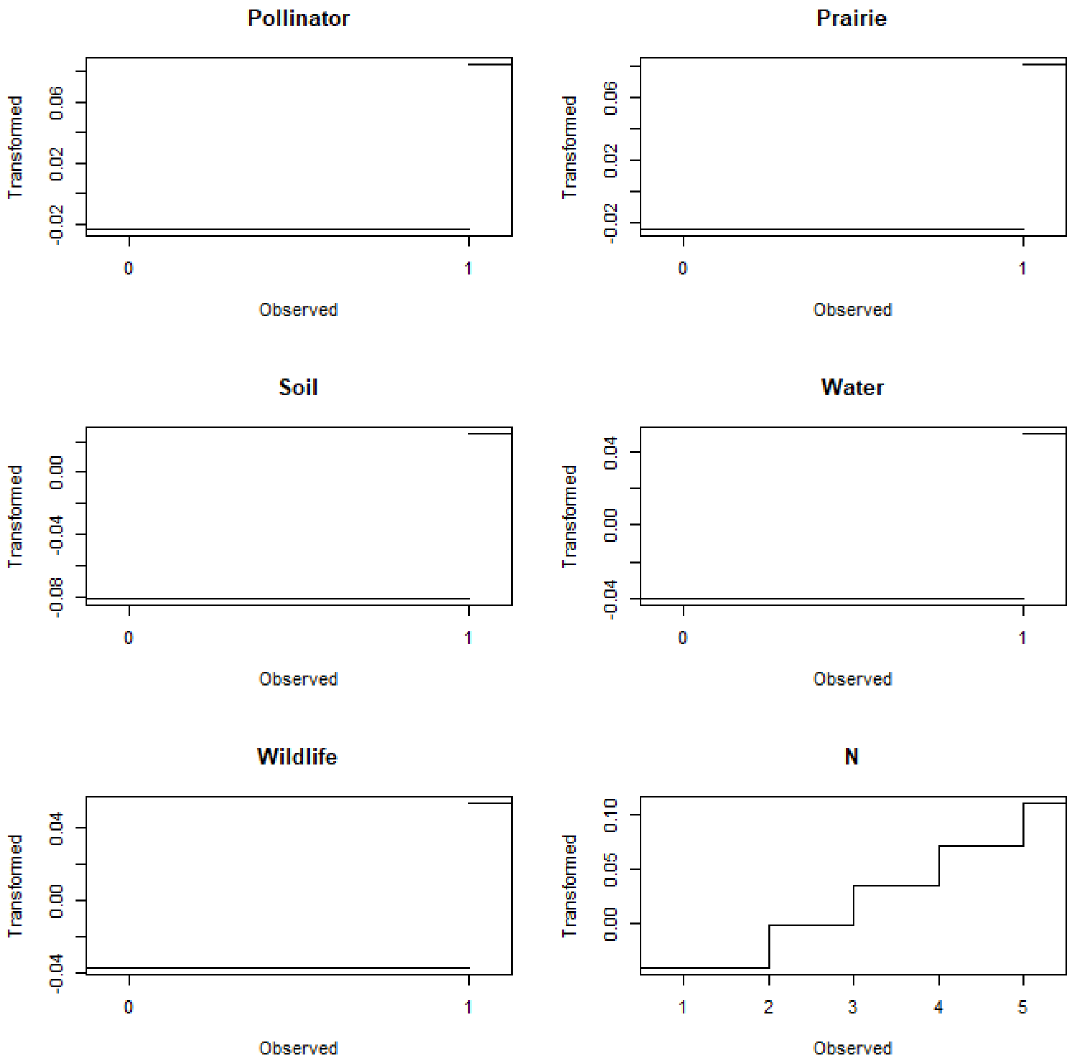


Figure A.1.10: Transformation plots for CATPCA in Method III.

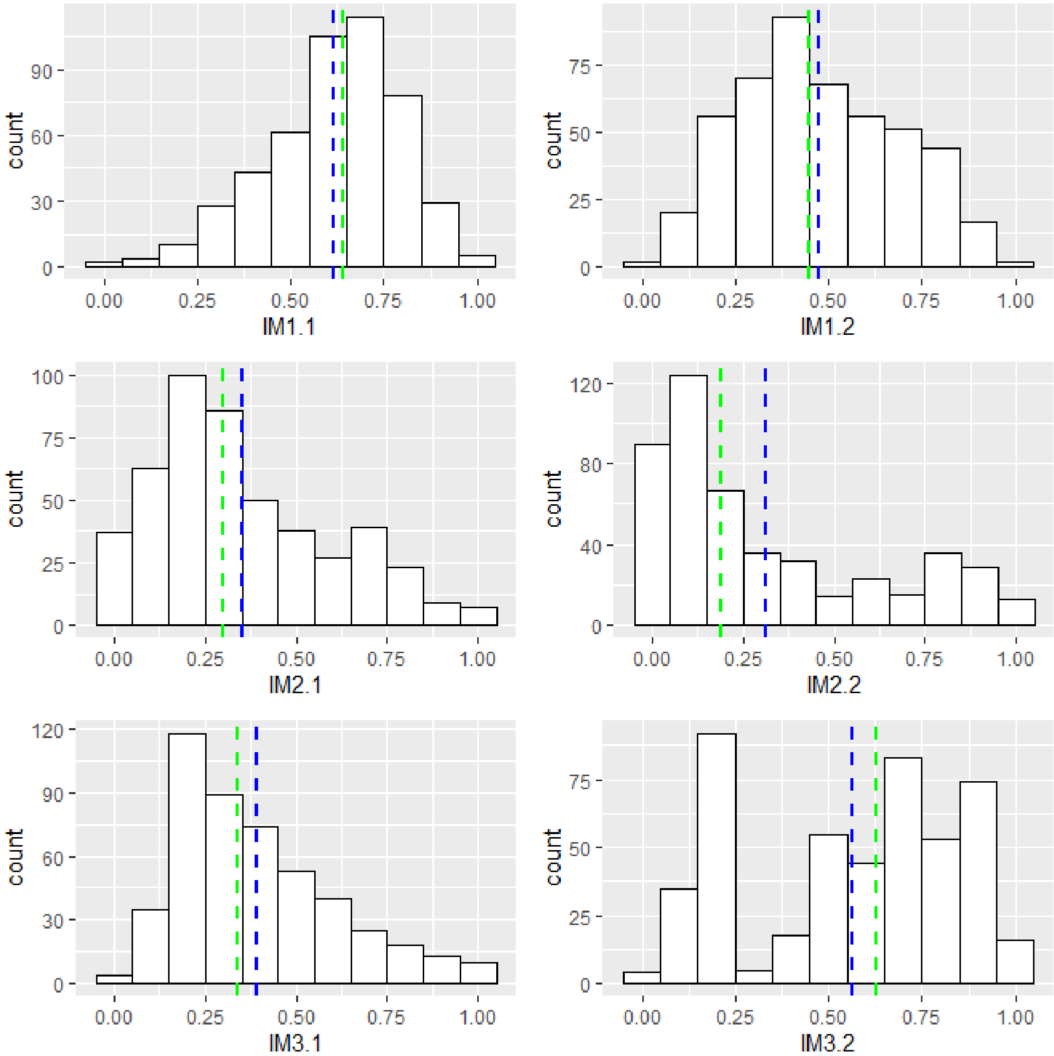


Figure A.1.11: Distribution for IM proxies.

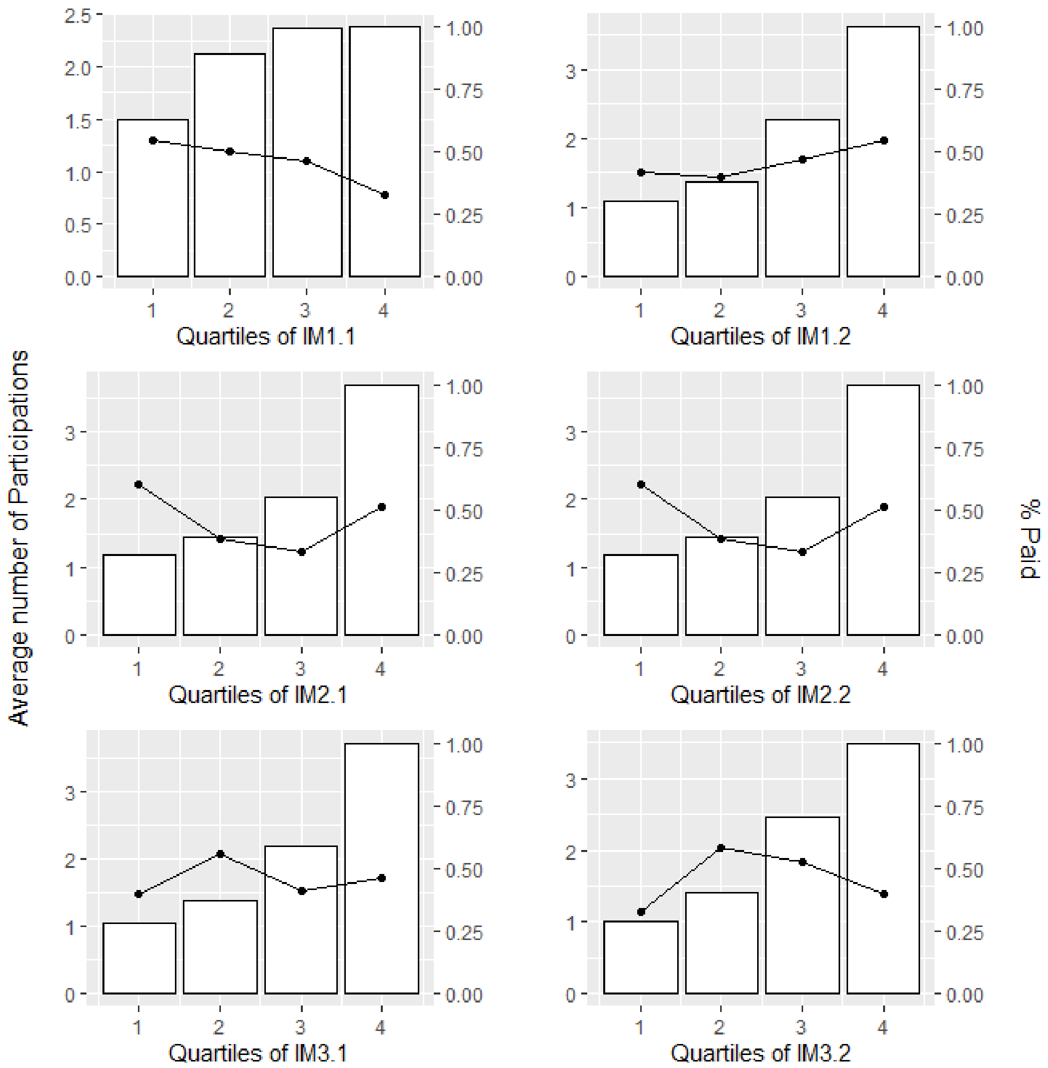


Figure A.1.12: Average number of programs farmers participate in and percentage of farmers paid by quartiles for all IM proxies.

A.2

	Dependent variable:
	Dummy variable for payments
Pollinator	0.576***
	(0.163)
Prairie	0.412**
	(0.161)
Soil	−0.002
	(0.157)
Water	0.188
	(0.131)
Wildlife	0.220
	(0.137)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	0.132***
	(0.046)
Two-year college/Technical/Vocational degree	−0.214
	(0.165)
Four-year college or more	0.025
	(0.147)
Age	−0.003
	(0.005)
Constant	−0.180
	(0.451)
Observations	479
Log Likelihood	-294.382
Akaike Inf. Crit.	658.764
Note:	*p<0.1; **p<0.05; ***p<0.01

Table A.2.1: Probit regression using dummy variable for payments as the dependent variable.

	D1	D2	D3	D4	D5
Pollinator	−0.627	0.031	0.232	−0.325	0.115
Prairie	−0.666	0.239	−0.021	−0.043	−0.009
Soil	0.038	−0.677	−0.005	0.453	0.391
Water	−0.502	−0.538	−0.298	0.083	−0.284
Wildlife	−0.674	0.218	0.040	0.230	−0.273
Age	−0.057	0.155	0.800	0.472	0.085
Education	−0.264	0.273	−0.371	0.142	0.743
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	0.270	−0.692	0.203	−0.211	−0.002
Payment Type	−0.401	−0.199	0.341	−0.555	0.330
N	−0.935	−0.287	−0.040	0.166	−0.064

Table A.2.2: Component Loadings for the first five principal components, D1 to D5, from CATPCA in Method II.

	D1	D2	D3
Pollinator	−0.622	0.181	0.667
Prarie	−0.647	0.315	0.021
Soil	−0.007	−0.867	0.268
Water	−0.569	−0.504	−0.410
Wildlife	−0.667	0.292	−0.330
N	−0.971	−0.225	0.024

Table A.2.3: Component loadings from CATPCA done in Method III.

	Dependent variable:	
	$E_i(\hat{U}_{ij})$	$\hat{U}_j$
	(1)	(2)
Annual Payment	0.301***	0.523***
	(0.051)	(0.091)
Onetime Payment	0.171**	0.386***
	(0.076)	(0.136)
Two-year college/Technical/Vocational degree	0.200***	0.262**
	(0.065)	(0.116)
Four-year college or more	0.160***	0.333***
	(0.058)	(0.103)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	−0.048***	−0.076**
	(0.018)	(0.032)
Age	0.0005	0.001
	(0.002)	(0.004)
Constant	−0.658***	−0.788**
	(0.174)	(0.311)
Observations	479	479
R^2	0.160	0.180
Adjusted R^2	0.102	0.123
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.2.4: Linear Regression where different functions of -D1, the first principal component, are used to approximate the dependent variable.

		Variables	#obs	Mean
1	What types of field crops were planted on this farm in 2018? What was the average yield of these planted crops?			
	i.	Corn	698	68.8%
		Corn Yield	431	185.37
	ii.	Soybean	698	61.9%
		Soybean Yield	394	52.24
	iii.	Sugarbeet	698	1.7%
		Sugarbeet Yield	13	17.08
	iv.	Wheat	698	7.4%
		Wheat Yield	49	55.51
2	Acres planted		634	448.48
3	Acres rented		617	213.80
4	Livestock on farm		656	50%
5	Male		653	90.5%
6	White		641	98.5%

Note: The unit of yield is bushels/acre except for sugarbeet which is measured in tons per acre.

Table A.2.5: Summary statistics

	Activities	Mean
Please indicate which of the management practices listed below were used on any of the farm's field crop acres in 2018?		
i	No or reduced tillage	0.529
ii	Planted cover crops	0.232
iii	Planted herbicide tolerant seed	0.641
iv	Planted seeds with pesticide treatment	0.453
v	Rotated crops	0.719
vi	Scouting for insect, weed or disease pests	0.586
vii	Soil testing to guide fertilizer application	0.627
viii	Spring fertilizer application	0.697
ix	Use of a post-harvest herbicide	0.102
x	Use of a foliar insecticide	0.196
xi	Use of a post-emergent herbicide	0.552
xii	Use of a pre-emergent herbicide	0.474
xiii	Use of a soil-applied insecticide	0.078
xiv	Use of tillage to control weed	0.347

Table A.2.6: Farm management practices of farmers.

	Conservation Practices	Mean
1.	What type of conservation practices do you currently participate in?	
i	Pollinator habitat	0.164
ii	Prairie wetland	0.179
iii	Soil conservation	0.587
iv	Water Quality	0.340
v	Wildlife habitat	0.316
2.	Have you received financial or other assistance for your conservation activities over the past 5 years?	
i	Annual payment	0.331
ii	Onetime payment	0.092
iii	Technical assistance	0.077
iv	Planning assistance	0.075
3.	Have you ever participated in any of the following USDA conservation programs?	
i	CREP	0.066
ii	CRP	0.410
iii	EQIP	0.175
iv	WRP	0.040
v	WHIP	0.020
vi	CSP	0.002
vii	ACEP	0.020

Table A.2.7: Conservation Practices of farmers.

	Mean Difficulty Level			
	Establishment Type			Preference based on the ease of use
Requirements	A	B	C	
Location Choice	2.451	2.380	2.414	-
Pre-planting	2.896	2.546	2.433	$C \succ B \succ A$
Planting	2.984	2.965	2.396	C
Watering	3.505	3.697	3.364	C
Mowing 1^{st} year	2.812	2.851	2.382	C
Mowing 2^{nd} year	2.714	2.675	2.369	C

Likert scale: “Not at all”-1,“Slightly”-2,“Somewhat”-3,“Very”-4,“Extremely”-5

Table A.2.8: Average difficulty level of requirements by the monarch habitat establishment types.

	Dependent variable:	
	Participation in monarch Conservation	
	(1)	(2)
\$value	0.001***	
	(0.0002)	
Paid	−0.417**	
	(0.191)	
\$100 per acre		−0.177
		(0.204)
\$250 per acre		−0.490**
		(0.207)
\$500 per acre		0.042
		(0.210)
\$1000 per acre		0.331†
		(0.212)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	−0.072	−0.079*
	(0.045)	(0.046)
Two-year college/Technical/Vocational degree	0.016	0.028
	(0.172)	(0.174)
Four year college or more	0.503***	0.496***
	(0.155)	(0.156)
Constant	−0.310	−0.259
	(0.484)	(0.488)
Observations	430	430
Log Likelihood	-267.958	-265.476
Akaike Inf. Crit.	605.916	604.953
Note:	†p<0.15; *p<0.1; **p<0.05; ***p<0.01	

Table A.2.9: Probit Regression with participation in the monarch conservation program as the dependent variable.

	Dependent variable:			
	Participation in Monarch conservation			
	(1.1)		(1.2)	
Paid	-0.420^{**}	-0.437^{**}	-0.433^{**}	-0.463^{**}
	(0.191)	(0.192)	(0.192)	(0.195)
IM	0.273	0.791	0.720^{**}	1.336^{***}
	(0.592)	(0.725)	(0.360)	(0.491)
\$value	0.001^{***}	0.002^{**}	0.001^{***}	0.002^{***}
	(0.0002)	(0.001)	(0.0002)	(0.0005)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	-0.060	-0.057	-0.048	-0.047
	(0.053)	(0.053)	(0.047)	(0.047)
Two-year college/Technical/ Vocational degree	-0.015	-0.014	-0.068	-0.052
	(0.185)	(0.185)	(0.178)	(0.178)
Four year college or more	0.502^{***}	0.492^{***}	0.465^{***}	0.442^{***}
	(0.155)	(0.156)	(0.157)	(0.158)
IM×\$value		-0.001		-0.002^{*}
		(0.001)		(0.001)
Constant	-0.445	-0.726	-0.507	-0.746
	(0.570)	(0.616)	(0.498)	(0.517)
Observations	430	430	430	430
Log Likelihood	-267.852	-267.056	-265.936	-264.156
Akaike Inf. Crit.	607.705	608.112	603.872	602.312
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.2.10: Probit regression with participation in the monarch habitat conservation program as the dependent variable and intrinsic motivation proxies 1.1 and 1.2.

	Intrinsic Motivation Proxy			
	(2.1)		(2.2)	
Paid	-0.438^{**}	-0.468^{**}	-0.441^{**}	-0.471^{**}
	(0.192)	(0.195)	(0.193)	(0.195)
IM	0.533^{*}	0.945^{**}	0.506^{**}	0.835^{**}
	(0.303)	(0.416)	(0.238)	(0.330)
\$value	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	(0.0002)	(0.0004)	(0.0002)	(0.0003)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	-0.054	-0.058	-0.051	-0.054
	(0.047)	(0.047)	(0.047)	(0.047)
Two-year college/Technical/	-0.010	-0.002	-0.014	0.001
Vocational degree	(0.174)	(0.174)	(0.174)	(0.175)
Four year college or more	0.431^{***}	0.414^{**}	0.423^{***}	0.410^{**}
	(0.161)	(0.161)	(0.160)	(0.160)
IM×\$value		$-0.001^{\dagger}$		$-0.001^{\dagger}$
		(0.001)		(0.001)
Constant	-0.242	-0.349	-0.191	-0.266
	(0.486)	(0.494)	(0.488)	(0.493)
Observations	430	430	430	430
Log Likelihood	-266.375	-265.247	-265.657	-264.558
Akaike Inf. Crit.	604.751	604.494	603.315	603.115
Note:	††p<0.20; †p<0.15; *p<0.1; **p<0.05; ***p<0.01			

Table A.2.11: Probit regression with participation in the monarch habitat Conservation program as the dependent variable and IM proxies 2.1 and 2.2.

	Intrinsic motivation proxy			
	IM3.1		IM3.2	
Paid	-0.441^{**}	-0.462^{**}	-0.447^{**}	-0.477^{**}
	(0.192)	(0.194)	(0.193)	(0.195)
IM	0.618^{**}	0.909^{**}	0.582^{**}	0.881^{**}
	(0.305)	(0.438)	(0.250)	(0.358)
$value	0.001^{***}	0.001^{**}	0.001^{***}	0.001^{**}
	(0.0002)	(0.0004)	(0.0002)	(0.0005)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	-0.061	-0.064	-0.059	-0.062
	(0.046)	(0.046)	(0.046)	(0.046)
Two-year college/Technical/	-0.016	-0.010	-0.039	-0.034
Vocational degree	(0.174)	(0.174)	(0.175)	(0.175)
Four year college or more	0.467^{***}	0.455^{***}	0.464^{***}	0.453^{***}
	(0.157)	(0.157)	(0.157)	(0.157)
IM×$value		-0.001		-0.001
		(0.001)		(0.001)
Constant	-0.416	-0.500	-0.501	-0.611
	(0.490)	(0.500)	(0.495)	(0.506)
Observations	430	430	430	430
Log Likelihood	-265.851	-265.408	-265.237	-264.559
Akaike Inf. Crit.	603.702	604.816	602.475	603.118

Table A.2.12: Probit Regression with participation in the monarch habitat conservation program as the dependent variable and IM proxies 3.1 and 3.2.

	Intrinsic Motivation Proxy					
	(1.1)	(1.2)	(2.1)	(2.2)	(3.1)	(3.2)
IM	0.461	0.834	−0.075	0.022	0.365	0.730
	(0.936)	(0.722)	(0.619)	(0.500)	(0.709)	(0.546)
Paid	−0.809	−0.876*	−1.051***	−0.871***	−0.787**	−0.621
	(0.683)	(0.478)	(0.335)	(0.274)	(0.390)	(0.441)
$value	0.002**	0.002***	0.002***	0.001***	0.001***	0.001**
	(0.001)	(0.001)	(0.0004)	(0.0003)	(0.0005)	(0.001)
$\log(\frac{\text{Farm size}}{\text{Median Farm size}})$	−0.054	−0.044	−0.055	−0.052	−0.063	−0.062
	(0.053)	(0.047)	(0.047)	(0.047)	(0.046)	(0.046)
Two-year college/ Technical/	−0.016	−0.062	−0.034	−0.030	−0.020	−0.036
Vocational degree	(0.185)	(0.179)	(0.176)	(0.176)	(0.175)	(0.176)
Four year college	0.493***	0.432***	0.397**	0.394**	0.445***	0.447***
or more	(0.156)	(0.158)	(0.161)	(0.161)	(0.157)	(0.158)
IM × Paid	0.608	0.881	1.695**	1.317**	0.859	0.268
	(1.068)	(0.932)	(0.793)	(0.634)	(0.898)	(0.735)
IM×$value	−0.002	−0.002**	−0.002**	−0.002**	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	−0.518	−0.487	0.046	0.026	−0.284	−0.518
	(0.719)	(0.585)	(0.527)	(0.513)	(0.547)	(0.564)
Observations	430	430	430	430	430	430
Log Likelihood	-266.891	-263.697	-262.902	-262.387	-264.943	-264.492
Akaike Inf. Crit.	609.782	603.395	601.805	600.774	605.887	604.984
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.2.13: Effect of interaction between IM proxy and getting paid on the probability of participation in the monarch conservation program.

Table A.2.14: Attributes of the monarch habitat conservation programs.

	1^{st} year establishment requirements			
	Attributes	Program 1	Program 2	Program 3
1.	Provided by the program for free	Milkweed seeds and peat pots	Milkweed root plugs	Seed mix with 5% native milkweed, 15% native grasses and 80% native wildflower
2.	Location	Has at least 6 hours of daylight and not treated by insecticides	Has at least 6 hours of daylight and not treated by insecticides	Has at least 6 hours of daylight and not treated by insecticides
3.	Pre-planting	Start milkweed seed in the peat pots 4 weeks before the last expected frost. Keep the seedlings moist, indoors and in sun until they are transplanted.	None	Apply any non-selective herbicide to the habitat location in the fall and again in the spring to clear competing vegetation.
4.	Planting	Plant individual seedlings 15 feet apart. Remove competing vegetation around seedlings in a 2-foot diameter circle.	Plant root plugs 4 to 6 inches deep and 15 feet apart. Remove competing vegetation around planted root plugs in a 2-foot diameter circle.	Broadcast seed mix at a rate of 50 pure live seed per square foot. Use roller, truck, or tractor to firm the seed into the soil.
5.	Watering	If rainfall isn't sufficient, water the milkweed.	If rainfall isn't sufficient, water the milkweed.	None
6.	Mowing	Mow monthly around milkweed to control competing weeds.	Mow monthly around milkweed to control competing weeds.	Mow monthly to control competing weeds.
	Year 2-5 Maintenance requirements			
	• Do not use insecticides or herbicides on the monarch habitat.			
	• Mow the monarch habitat only in the fall of year 3.			

B.1

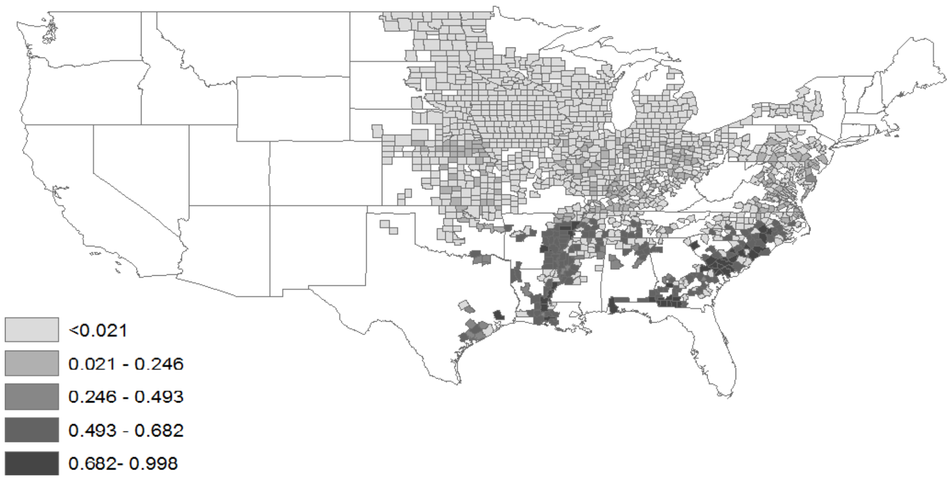
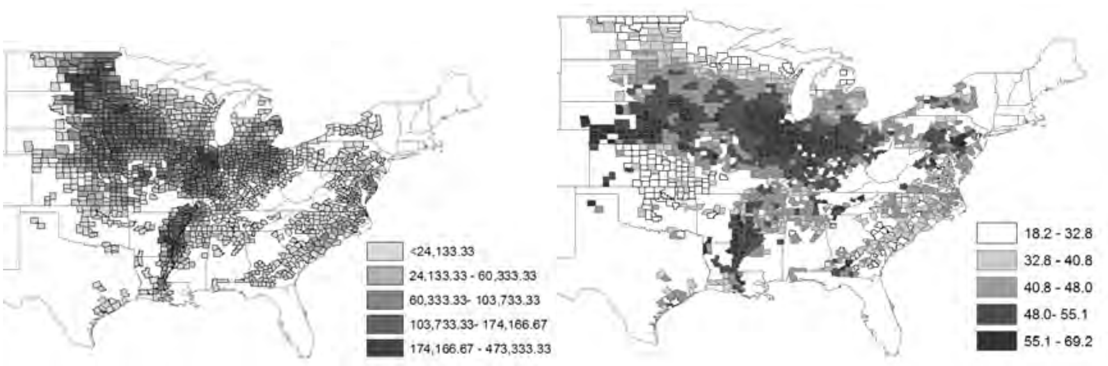


Figure B.1.1: Probability of infection.

Parameter	Value	Reference
Scouting Costs C_{sc}	\$8.68/Acre	[128, 127][1]
Cost of Curative Fungicide C_c	\$17.86/Acre	"
Cost of Preventative Fungicide treatment C_p	\$33.15/Acre	"
Yield loss with no treatment λ_n	25%	"
Yield loss with Curative Fungicide treatment λ_r	7%	"
Discount Factor δ	0.94	"

[1] C_{sc}, C_c and C_p reported here are inflation adjusted.

Table B.1.1: Baseline parameters for the farmer decision model.



(a) (b)

Figure B.1.2: (a) Soybean acreage and (b) yield (bu/acre) in 2014.

Variable	Observation level	Average	Standard Dev	Max	Min	Total observations
Price ($/bu)	State	10.096	0.356	11	9.37	1360
Yield (bu/acre)	County	45.664	9.365	69.2	18.2	1360
Acreage ('000 acres)	County	49.275	50.454	473.33	0.3	1360
Prior	County	0.096	0.21	0.99	0	1360

Table B.1.2: Summary Statistics

B.2

Let i be defined as the occurrence of infection in the field and the probability of infection $(i = 1)$ is ϕ^f. Similarly, j denotes infection in the sentinel plot, and ϕ^s the probability that j=1. Hence

$$
\begin{aligned}
E(i) &= \phi^f && (16)\\
E(j) &= \phi^s && (17)\\
Var(i) &= \phi^f(1-\phi^f) && (18)\\
Var(j) &= \phi^s(1-\phi^s) && (19)\\
E(ij) &= \phi_{11} && (20)
\end{aligned}
$$

Therefore, correlation ρ can be expressed as

$$
\begin{aligned}
\rho &= \frac{Cov(i,j)}{\sqrt{Var(i)Var(j)}}\\
&= \frac{E(ij)-E(i)E(j)}{\sqrt{Var(i)Var(j)}}\\
&= \frac{\phi_{11}-\phi^s\phi^f}{\sqrt{\phi^s(1-\phi^s)\phi^f(1-\phi^f)}}\\
\implies \phi_{11} &= \phi^s\phi^f+\rho\sqrt{\phi^s(1-\phi^s)\phi^f(1-\phi^f)}. && (21)
\end{aligned}
$$

Also,

$$
\begin{aligned}
\phi_{11}+\phi_{10} &= \phi^f && (22)\\
\phi_{11}+\phi_{01} &= \phi^s && (23)\\
\phi_{00}+\phi_{11}+\phi_{10}+\phi_{01} &= 1. && (24)
\end{aligned}
$$

Using the above equations, we get:

$$
\begin{aligned}
\phi_{01} &= (1-\phi^f)\phi^s-\rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} && (25)\\
\phi_{10} &= \phi^f(1-\phi^s)-\rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)} && (26)\\
\phi_{00} &= (1-\phi^f)(1-\phi^s)+\rho\sqrt{\phi^f(1-\phi^f)\phi^s(1-\phi^s)}. && (27)
\end{aligned}
$$

C.1

	Variable	Transition rule	References
1.	Nymphs	$N_t^{ij\tilde{a}} = \lambda N_{t-1}^{i}[I(j=w)p(j=w\|t,i)$	
		$+I(j=\tilde{w})(1-p(j=w\|t,i))]$	(4.4),(4.7)
2.	Total Adults	$N_t^{ija} = N_t^{ija}(1-E_t^{ij}) + \sum_z E_t^{zji}N_t^{zj}$	
	i. Indigenous Adults	$N_t^{ija} = (1-\delta F_{t-1}^{i})N_{t-1}^{ija}$	(4.4)
	ii. Emigrants	$N_t^{ija}E_t^{ij}$	(4.5)
	iii. Immigrants	$\sum_z E_t^{zji}N_t^{zj}$	(4.5), (4.6)
3.	Total aphid j	$N_t^{ij} = N_t^{ij\tilde{a}} + N_t^{ija}$	
4.	Total aphids	$N_t^{i} = N_t^{i\tilde{w}} + N_t^{iw}$	

Table C.1.1: Transition rules for the state variables of a cell i at time t.

[1] Superscripts $\{i,j,l\}$ in N_t^{ijl} represent host cell location, morph and age group respectively. Therefore, $i \in \{1,2,..M_1M_2\}$, $j \in \{w,\tilde{w}\}$ and $l \in \{a,\tilde{a}\}$.
[2] $I(j=w)$ is an indicator variable which equals 1 if $j=w$ and 0 otherwise. On the other hand, $I(j=\tilde{w}) = 1 - I(j=w)$.
[3] Superscripts $\{i,j,h\}$ on variable E_t^{ijh} denote host cell i, morph j and destination cell h respectively. Cell h is within the neighborhood of cell i using Equation 4.6. Therefore, in 2(iii) and (iv), $i \in \mathcal{N}(j|z)\ \forall j$ and $i \neq z$.

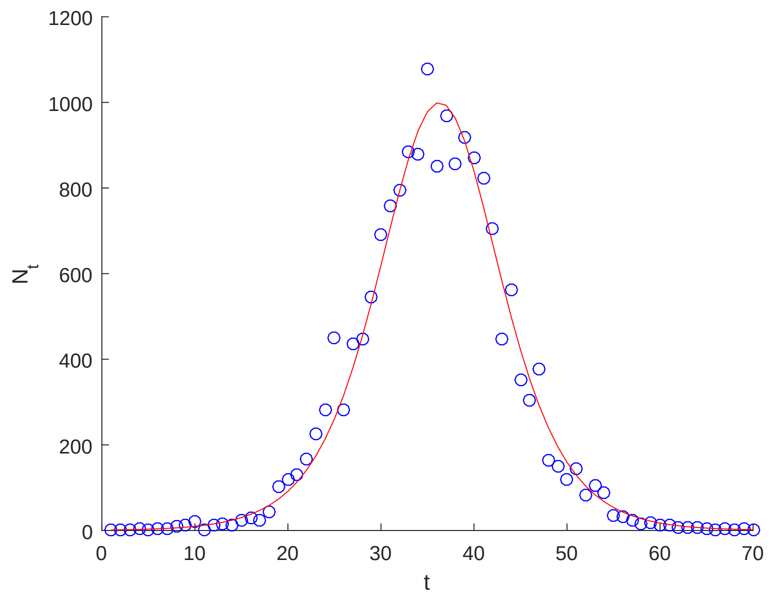


Figure C.1.1: Population N_t as a function of t when $N_{max} \sim N(1000, 100)$, $T_{max} \sim N(36.29, 0.42)$ and $b \sim N(0.228, 0.23)$.

P(Type I Error)	P(Type II Error)					
	0	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
0	1.6	1.6	1.6	1.6	1.6	1.6
0.1	1.6	1.6	1.6	1.6	1.6	1.6
0.2	1.6	1.6	1.6	1.6	1.6	1.7
0.3	1.6	1.6	1.6	1.6	1.6	2.1
0.4	1.6	1.6	1.6	1.7	2.2	6.4
0.5	1.6	1.6	1.7	2.1	6.4	6.2

Table C.1.2: Percentage yield loss as a function of the probabilities of type 1 and 2 error.

Parameter	Symbol	Value/Functional form	Reference
Maximum days	t_{max}	50	[111]
Rows & columns in field	M_1 & M_2	10	Assumption
Price of soybean ($/ton)	P	220.46	[122]
Potential Output (ton/ha)	Y_P	4.04	[122]
Cost of monitoring ($/ha)	C^k	9.88 $\forall\ k = m, u$	[74]
Treatment Cost ($/ha)	C_I	35.82	[74]
Yield loss function parameters	α_1	4	[122]
	α_2	10	Assumption
	α_3	0.00013	”
	α_4	6	”
Efficacy of Insecticide	θ	0.99	[157]
Plants sampled during manual scouting		20	[111]
Mean of aphid count at t=0	m	5	Assumption
Variance of aphid count at t=0	ν	$9.152 \times m^{1.543}$	[63]
Birth and Death rate parameters	n_{max}	1,000±100	[82]
	T_{max}	36.29±0.42	”
	b	0.228±0.23	”
Radius of neighborhood of alates		0.5km	[158]
Radius of neighborhood of aptera		10 inches	Assumption
Probability that alate migrates at time t	E_t^{iw}	0.04 $\forall\ i$	[32]
Tally threshold of UAV	ϕ	250 aphids per plant	[98]

Table C.1.3: Parameters and their values.

	P(Type II Error)						
P(Type I Error)	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**Average**
0	785.18	785.18	785.18	785.18	785.18	785.23	785.19
0.1	785.18	785.18	785.18	785.18	785.23	785.19	785.19
0.2	785.18	785.18	785.18	785.18	784.98	784.59	785.05
0.3	785.18	785.18	785.18	785.18	784.85	781.18	784.46
0.4	785.18	785.18	784.87	784.01	780.71	777.55	782.92
0.5	785.18	784.97	784.35	781.71	777.96	778.15	782.05
Average	785.18	785.15	784.99	784.41	783.15	781.98	784.14

Table C.1.4: Optimal expected profit from UAV based scouting, $E(\pi_u^*)$, as a function of the probabilities of Type 1 and 2 error in \$ per hectare.

	P(Type II Error)					
P(Type I Error)	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
0	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0.02,0.98,0,0]
0.1	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0.02,0.98,0,0]	[0.04,0.96,0,0]
0.2	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0.02,0.98,0,0]	[0.98,0.02,0,0]	[0.04,0.9,0.06,0]
0.3	[0,1,0,0]	[0,1,0,0]	[0,1,0,0]	[0.02,0.98,0,0]	[0.02,0.94,0.04,0]	[0.6,0.38,0.02]
0.4	[0,1,0,0]	[0,1,0,0]	[0,0.98,0.02,0]	[0.02,0.86,0.12,0]	[0.04,0.62,0.26,0.06]	[0,0.02,0,0]
0.5	[0,1,0,0]	[0,0.98,0.02,0]	[0,0.92,0.08,0]	[0.02,0.64,0.3,0.02]	[0,0.02,0,0]	[0.02,0,0.02,0.02,0,0]

Note: $E(I_7^u)]$ is always 0. $E(I_1^u)]$ and $E(I_2^u)]$ are also 0 except when probabilities of type 1 and 2 errors equal 0.5. Therefore, we report $[E(I_3^u), .., E(I_6^u)]$ for every combination of probabilities other than when they equal 0.5 for which $[E(I_1^u), .., E(I_6^u)]$ is reported.

Table C.1.5: The likelihood of treatment as a function of the probabilities of type 1 and 2 error.

C.2

To estimate $\{\mu, \sigma\}$, we use the mean m and variance ν of N_0. $\{\mu, \sigma\}$ are related to $\{m, \nu\}$ in the following way [93]:

$$\mu = log\left(\frac{m^2}{\sqrt{\nu + m^2}}\right) \tag{28a}$$

$$\sigma = \sqrt{log(\frac{\nu}{m^2 + 1})} \tag{28b}$$

In order to calculate ν, Taylor's power law is used. Taylor's power law relates the variance of the number of individuals of a species per unit area of habitat to the corresponding mean by a power law relationship. Hodgson *et al.* (2004, [63]) estimate the parameters of Taylor's power law equation for soybean aphid. Using their estimates, we know that

$$\nu = 9.15 \times m^{1.543}. \tag{29}$$

We estimate $\{\mu, \sigma\}$ using equations (28a), (28b) and (29).Note that for $\sigma > 0$, $m^2 + 1 > 9.5m^{1.5}$ using (28a), (28b) and Taylor's law.