Three Essays on Innovation:

Optimal Licensing Strategies, New Variety Adoption, and Consumer Preference in a Peer Network

by

Di Fang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved September 2015 by the
Graduate Supervisory Committee:

Timothy Richards, Chair
Ruth Bolton
Carola Grebitus
Mark Manfredo

ARIZONA STATE UNIVERSITY

December 2015

ABSTRACT

It is well understood that innovation drives productivity growth in agriculture. Innovation, however, is a process that involves activities distributed throughout the supply chain. In this dissertation I investigate three topics that are at the core of the distribution and diffusion of innovation: optimal licensing of university-based inventions, new variety adoption among farmers, and consumers' choice of new products within a social network environment.

University researchers assume an important role in innovation, particularly as a result of the Bayh-Dole Act, which allowed universities to license inventions funded by federal research dollars, to private industry. Aligning the incentives to innovate at the university level with the incentives to adopt downstream, I show that non-exclusive licensing is preferred under both fixed fee and royalty licensing. Finding support for non-exclusive licensing is important as it provides evidence that the concept underlying the Bayh-Dole Act has economic merit, namely that the goals of university-based researchers are consistent with those of society, and taxpayers, in general.

After licensing, new products enter the diffusion process. Using a case study of small holders in Mozambique, I observe substantial geographic clustering of new-variety adoption decisions. Controlling for the other potential factors, I find that information diffusion through space is largely responsible for variation in adoption. As predicted by a social learning model, spatial effects are not based on geographic distance, but rather on neighbor-relationships that follow from information exchange. My findings are consistent with others who find information to be the primary barrier to adoption, and means that adoption can be accelerated by improving information exchange among farmers.

Ultimately, innovation is only useful when adopted by end consumers. Consumers' choices of new products are determined by many factors such as personal preferences, the attributes of the products, and more importantly, peer recommendations. My experimental data shows that peers are indeed important, but "weak ties" or information from friends-of-friends is more important than close friends. Further, others regarded as experts in the subject matter exert the strongest influence on peer choices.

To my parents, two friends, and two dogs.

Without their love and patience,

I would have finished this dissertation sooner.

ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to those who helped me along the journey. First and foremost, I would like to thank my advisor, Dr. Timothy Richards, with who I worked closely for six years. His intelligence, patience and consistent support are the reasons I can make it through the particular hardships of achieving a Ph.D. degree. I benefited especially from his insightful guidance over the course of my dissertation. His elaborate and thoughtful comments inspired this project while his efficient feedbacks helped me finish it.

I am very grateful to the rest of my committee members: Dr. Ruth Bolton, Dr. Carola Grebitus, and Dr. Mark Manfredo for their guidance, support and encouragements. Dr. Bolton has been a role model and a great friend to me-- her insightful perspectives have helped me immensely. I also appreciate the valuable advices from Dr. Troy Schmitz. I thank all the faculty and staff of Morrison School of Agribusiness for their support. I would also like to thank my fellow Ph.D. students for their companionship--because of them I never felt alone.

The generous support from the Agriculture and Food Research Initiative (AFRI) - National Institute for Food and Agriculture (NIFA), USDA is gratefully acknowledged. I would also like to thank the Michigan State University for the data they provided. Information on availability and access to the Data are available at: http://fsg.afre.msu.edu/Mozambique/survey/index.htm.

Most importantly, I want to thank my family, friends and my dogs, you are the reason I can pursue this career and continue to do so. I owe every joy of satisfaction and future success to you. I thank my father, Jiyou Fang, who is behind every one of my

decisions and more, and my mother, Yinghua Xu, who shows me there is always home.

"And I tell you, ask, and it will be given to you; seek, and you will find; knock, and it will be opened to you. For everyone who asks receives, and the one who seeks finds, and to the one who knocks it will be opened."
-- Luke 11:9-10

TABLE OF CONTENTS

Page

LIST OF TABLES ... ix

LIST OF FIGURES ... x

CHAPTER

1. INTRODUCTION ... 1

2. ESSAY 1: FEES VERSUS ROYALTIES IN AGRICULTRAL PATENTS ... 13

 2.1 Introduction ... 13

 2.2 Model ... 16

 2.3 Propositions Regarding the Innovator's Profit ... 19

 2.4 Conclusion ... 36

 2.5 Discussion. ... 40

3. ESSAY 2:VARIETY ADOPTION AMONG FARMERS IN A SOCIAL NETWOR 42

 3.1 Introduction ... 42

 3.2 Theoretical Model ... 48

 3.3 Hypotheses ... 57

 3.4 Empirical Study: New Maize Variety Adoption in Mozambique ... 64

  3.4.1 Data ... 65

  3.4.2 Model ... 68

  3.4.3 Estimation ... 77

  3.4.4 Results ... 80

CHAPTER Page

3.5 Conclusion 85

4. ESSAY 3: MODELING PRODUCT CHOICES IN A PEER NETWORK........ 89

4.1. Introduction........ 89

4.2 Background........ 92

4.3 Conceptual Model of Endogenous Peer Effects 99

4.4. Experimental Procedure........ 103

4.5 Econometric Model of Preference Revision........ 115

4.5.1 Choice modeling........ 116

4.5.2. Spatial Models 122

4.5.3 Estimation 129

4.6 Results........ 133

4.6.1 Multinomial Logit Results........ 134

4.6.2 RPL Results 136

4.6.3 Results of Spatial Models 141

4.7 Conclusion 147

5. CONCLUSIONS AND IMPLICATIONS........ 150

REFERENCES 160

APPENDIX

A CALCULATING INNOVATOR'S PROFIT........ 167

B BAYESIAN UPDATING 173

C EXPERIMENT INSTRUCTIONS........ 176

LIST OF TABLES

Table Page

2.1: Comparison of Innovator Profits and the Extent of Innovation................................ 39

2.2: Comparison of Profits and Horizontal Differentiation 40

3.1: The Number of Adopters in the Sampled Provinces 67

3.2 Selected Variables and Descriptive Statistics. 74

3.3 Diagnostics for Spatial Dependence Test Against a Classic Probit........................... 82

3.4 Spatial Models Using Different Weights. 83

4.1 Attributes and Levels. 107

4.2 Example Choice Set 108

4.3 Descriptive Statistics for the Sample 114

4.4 List of Variables... 118

4.5 Variables Included in the Model 127

4.6 Estimated Parameters for MNL And RPL. 134

4.7 Subject-Specific Marginal WTP by Attribute.. 140

4.8 Results of Spatial Models. 146

LIST OF FIGURES

Figure Page

2.1 Innovator Profit Under Exclusive Fixed-Fee Licensing 23

2.2 Innovator Profit Under Non-Exclusive Fixed-Fee Licensing 25

2.3 Innovator Profit Under Exclusive Royalty Licensing 28

2.4 Innovator Profit Under Non-Exclusive Royalty Licensing 30

2.5 Innovator Profit Under Exclusive Two-Part Tariff Licensing 34

2.6 Innovator Profit Under Non-Exclusive Two-Part Tariff Licensing 36

3.1 Sampled Household in Sofala and Manica with Adopters (Red) and Non-Adopters (Black) 43

CHAPTER 1. INTRODUCTION

Innovation is key to not only agricultural but all economic progress. Productivity improvement from high yield varieties and more efficient machinery freed farmers from the land, which, in turn, enabled the modern consumer economy. For example, the Green Revolution of the 1940s to the 1960s benefited not only developing countries, but also the world, through lower prices, high yields, and enhanced productivity due largely to improved crop varieties (Evenson and Gollin 2003). Although many innovations arise from serendipity, truly transformative technologies emerge from an innovation process, a process that mirrors the agricultural value chain from the supply of innovation at research universities, through adoption by producers, and acceptance by consumers. In this dissertation, I study issues at the core of each stage of the innovation process: Licensing patents from research universities, adoption of new varieties through learning process, and how social networks affect the adoption of new consumer products.

Every year, universities license creates approximately $30 billion profit through patents to industry. Enabled by the Bayh-Dole Act (1980), which granted universities the right to profit from patents on federally-funded research, and incentivized through the reduction in public funding for university budgets, universities actively market innovations that emerge from their labs through licenses managed by university-based technology transfer offices (TTOs). However, despite the rising importance of licensing revenue to university funding, patent licenses are not marketed optimally by TTOs. Pricing innovations correctly not only aligns the incentives facing researchers and their university employers but helps ensure that a more sustainable stream of innovations enters the agricultural supply chain. In this dissertation, I resolve the seeming paradox that arises through the current system of patent licensing – a process with the ability and need to be self-supporting appears to be neither, in spite of the wealth of research on patent licensing schemes.[1]

Recognizing university-based intellectual property (IP) is necessary to protect the rights of the innovators and to motivate further breakthroughs. While the Bayh-Dole Act seemingly entrenched the right of universities to freely market their IP, this right has recently been challenged in court. Notably, the University of Minnesota licensed its SweeTango™ apple to a single group of apple growers -- 45 of them, mainly in the states of Washington, Michigan, and New York -- the Next Big Thing cooperative. SweeTango™ apples can only be grown by members of this cooperative. At issue is whether a group of growers should monopolize the right to utilize a university-based research, and whether the university, and by extension society, benefits from such exclusive licensing. To examine this matter, I compare different licensing scenarios to compare the welfare of both the patent holder and the patented growers (firms). Ultimately, my goal is to determine the licensing strategy that is in the best interests of university administrators. Robust licensing strategies help protect the profit of the innovator and regulate the market, therefore, finding the optimal licensing strategy is crucial.

[1] Optimality is defined, as in the literature, as the difference between license revenues and the cost of innovation. Although our analysis concerns university research activities, and universities are expected to conduct basic research in the public interest, our profit-maximization assumption reflects the observed activities of university TTO offices. Namely, as Bulut and Moschini (2009) note, "Quite clearly, when it comes to patenting and licensing, universities are likely to behave based on their self-interest rather than the public interest" (p. 124). Resolving the debate as to whether universities should maximize the returns to their research investment is left for either political or legal discussion.

Licensors generally pay for the rights to use a patent in one of three ways: through fixed fees, royalties, or a combination of fees and royalties in a two-part tariff scheme, each marketed through either exclusive or non-exclusive contracts. Optimal contracts, or those that maximize university revenue, depend critically on the nature of competition among downstream producers, or the firms that buy and use the new technology. While license design for cost-reducing innovations is relatively well understood (Sen and Tauman 2007), there is relatively little research on quality-improving innovations (Bousquet et al. 1998; Sen and Tauman 2007; Stamatopoulos and Tauman 2008; Li and Wang 2010).

In agriculture, in general, and in the fruit and vegetable sector more specifically, however, a growing number of innovations seek to improve either nutritional, taste, ethical, or other demand-side attributes. Unlike cost-reducing innovations, which do not alter downstream demands, quality improving innovations are important in that they vertically differentiate old and new products, leading to greater willingness to pay for consumers. In the case of the SweeTango™ apple described above, the eating qualities of the apple were thought to be a substantial improvement over existing varieties. Because there are many potential downstream competitors willing to license a patent on a new farm product, I consider an oligopolistic market where downstream firms compete in price and the upstream innovator holds a quality-improving technology that may vertically-differentiate old and new products. In other words, I study the optimal design of patent-licensing contracts in the context of a demand-side agricultural product innovation.

The innovator is assumed to be an outsider, or an entity that does not also produce the commodity itself, that sells licenses to its innovations through either exclusive or non-exclusive licensing. Exclusive licensing assumes sole ownership of the patent and creates asymmetrical demands for new and traditional products. Non-exclusive licensing, on the other hand, allows sharing of the patent and creates a market of only new products.

My findings differ from previous research in a fundamental way. First, I find that non-exclusive licensing is preferred under both a fixed fee and a royalty. This is important, because the SweeTango case, as just one example, revolved around whether the innovator should use exclusive or non-exclusive contracts. By showing that university benefits more with non-exclusive licensing, my findings are consistent with arguments made by the growers who were excluded from growing the SweeTango who argue that a group of growers should not be able to monopolize production. Indeed, this result also supports the broader mission of land grant universities in general as I show that publicly funded research is most efficient when made available to the largest number of growers possible.

Optimal licensing contracts are driven primarily by strategic considerations between the licensor and licensee in that, by offering a royalty-based contract, the innovator's profit is linked to the output of the licensing firms. On the other hand, offering a fixed-fee strategy allows the innovator to soften competition between the two downstream firms. Therefore, the innovator has some control over the market by choosing a licensing strategy: Licensing through a royalty will create more intense competition between downstream competitors while licensing by fixed fee will relax competition, and raise prices downstream.

Of course, the nature of the innovation is critical as well. I find that the innovator's license revenue depends on the magnitude of the innovation – defined as the difference between the new and existing products -- and the degree of substitutability between the products. Generally speaking, the more drastic the innovation, the greater the license revenue to the innovator. Because innovations create value when they sharpen the extent of differentiation between old and new products, there is a greater reward to the innovator's institution if the innovation is substantial, and firms can still differentiate downstream. For TTO managers, this finding suggests that research offices should foster an environment of risk taking if they want to maximize returns from their portfolio of research.

When all strategic aspects of downstream innovations are properly considered, the fundamental conflict between university and social objectives engendered by Bayh-Dole suggested by previous research (Folli-Oller and Sandonis 2005) disappears and instead suggests that the players' goals are aligned at a very basic level. The Next Best Thing case revolved around the exclusivity of access to the University of Minnesota's new apple. While the prior literature on patent licensing supported the case made by the University, my results suggest that the plaintiffs had a strong case: The objectives of the University, and Minnesota taxpayers more generally, are better served through a system of non-exclusive licensing.

Patent licensing is the first step of introducing an innovation; however, whether innovations will prove to be profitable over the long term ultimately determines whether they will be adopted, and lead to fundamental improvements in efficiency. For this reason, learning and imitation are central to adoption (Warner 1974). Before deciding whether to adopt, farmers seek information on the costs and benefits of the innovation from their own and other users' experiences. As they gather more information, they are able to increase their knowledge about the overall attractiveness of the innovation and reduce the uncertainty associated with its potential benefits. For this reason, any delay in adoption that is designed to allow for the acquisition of more information is inherently strategic as producers wait to learn from others' experiences (Besley and Case 1997).

At each point in time, strategic delay is manifest in an observed geographic clustering of adoption in a space that consists of farmers arrayed in a scattered social-spatial network. There is empirical evidence that the incentives to adopt depend on how many others have adopted, rising quickly and then falling in a quadratic pattern (Bandiera and Rasul, 2006), but no evidence that this is due to a social-learning mechanism.

Information is not complete in developing countries due to constraints placed by public transportation, government regulations, and the lack of social media. Therefore, incomplete information is a major barrier in variety adoption. Particularly in the context of rural areas of developing countries, infrastructure such as roads and markets are under-developed, which imposes constraints on the ability of farmers to observe their neighbors. Information comes from two sources: A farmer's own experience and from observing others. As the number of adopters rises in a network, expansion creates larger information externalities, providing non-adopters with more information (Banerjee 1992; Caplin and Leahy 1998; and Chamley and Gale 1994). Having connections with others helps a farmer update his knowledge about the new variety. However, direct links are costly (Bala and Goyal 2000) because of the effort required to form and maintain the connections.

Existing theoretical models of adoption, however, ignore the cost of observing information. Assuming away the cost of becoming informed changes the implications of the model in fundamental ways. For example, if information can be acquired costlessly, then farmers will choose to wait forever due to the positive externalities generated by learning (Zhao 2006). The outcomes of a costly-learning model are fundamentally different in that the relevant pure strategy to the adoption game with imperfect information will be altered due to the cost structure. More specifically, a farmer will need to consider both the network externalities as well as the cost of maintaining a network when making adoption decisions. While a model with costly information acquisition predicts strategic delay, costless information results in either immediate adoption or a bifurcated market in which some farmers adopt, and others never do. Both logic and observation suggest that the former is a more reasonable description of reality.

I assume farmers learn about new varieties through a pure strategy Bayesian equilibrium that includes both information exchange and costly network formation. I use this model to generate hypotheses regarding the effect of learning, extension services, size of a farmer's network, and the cost of maintaining a network on the speed of adoption. In each case, the hypothesis that results presents an empirical puzzle that must be answered by taking all effects into account at once.

My empirical analysis describes household-level variety adoption decisions made by farmers in Mozambique. Framed in a spatial econometric environment, I model each farmer's adoption decision in the context of his social network, using the distance to each other farmer as the location in network space. Specifically, the distance between one farmer and another forms a social "weight" that moderates the effect of social learning.

I find that there is significant, positive spatial correlation in adoption decisions among neighboring farmers, In other words, if one farmer adopts, a neighbor is more likely to adopt as well. Clearly, social learning is not only important to my sample farmers, but a dominant source of information. Education, on the other hand, is an important catalyst for adoption. Training is found to be an important parameter improved maize adoption, indicating that hands-on demonstration should be encouraged. Moreover, when farmers produce for sale instead of own consumption, they are more likely to form larger networks for cost sharing.

In this model, I compare three types of social proximity defined using a “nearest neighbor” metric, an arc-distance matrix, and a contiguity matrix. Contiguity matrices capture relationships among farms that share common boundaries, while an arc-distance matrix measures the geographical distance between two households. Finally, a nearest neighbor matrix measures immediate neighbors. Allowing for heterogeneity among sample observations, I find that although immediate neighbors have a positive influence on the probability of adopting a new variety, social learning is not constrained by distance. Furthermore, farmers extend their network to include more households for the purpose of information sharing.

In the Mozambique example, I implicitly assume consumer preferences for new varieties are exogenously determined. It is more likely, though, that consumer preferences and grower adoption are simultaneously determined in real-world environments. As a way of thinking about innovation creation, adoption, and diffusion as a process, my third essay examines how new products are accepted by consumers in the downstream-market.

An emerging literature, both in economics and marketing, shows that consumer choices are influenced not only by their own tastes, but the choices made by others. People who identify with a group often adopt similar tastes as the group (Case 1991; Yang and Allenby, 2003; Zimmerman 2003; De Giorgi, Pellizzari, and Redadelli 2010; Kuhn et. al. 2010; Richards, Hamilton and Allender 2014). Although the mechanisms differ among studies in this literature, the common element points to the conclusion that, when studying consumer choices, it is impossible to ignore the interdependence of preferences in a social network. In the third essay, I focus on a certain social interdependence—peer effects. Peer effects refer to that part of behavior among individuals that arises purely from the influence of others, whether through imitation, learning, or aversion.

Identifying peer effects empirically is problematic because observational data do not allow us to easily separate the influence of peer behavior on the behavior of others from that of observed and unobserved factors. Empirically, individuals in groups tend to behave in similar ways for three primary reasons: endogenous effects, contextual effects, and correlated effects (Manski, 1993; Brock and Durlauf, 2001; Mofitt, 2001; Soetevent, 2006). Endogenous effects refer to the true peer effect in that a consumer’s choices are influence by others, while contextual effects refer to the fact that people who come from the same background (such as education, income, etc.) tend to make similar decisions regarding product choices, independent of the peer influence, and unobserved, correlated effects are those who cause people to choose similarly but unseen by the obvious mechanisms. In this study, I focus on identifying endogenous peer effects by controlling for the other factors that can otherwise confound identification.

Identification relies on either clear randomization or clever econometric modeling, but preferably a combination of the two (Narayan, Rao and Sanders 2011; Yang and Allenby 2007; Richards, Hamilton and Allender 2014). Ideally, membership in the social network is completely random – as in the job assignments of Duflo and Saez (2002) – so that any peer effects that emerge can be due only to learning from others. However, purely random samples are rare, even in lab situations, so econometric modeling can help.

I use a two-stage lab experiment to gather data that describes subjects' adoption decisions as a function of peer choices. As in other peer-influence studies (Narayan, Rao, and Saunders 2011), the experiment consists of two stages: The first stage measures initial preferences while the second stage measures peer-based preference revisions. Substantial preference revision is evident from simple examination of the data, while econometric estimates reveal statistically significant peer effects.

I estimate the extent of preference revision using a spatial econometric model, where space is defined in terms of subjects' social relationships with other subjects. In this model, I have more flexibility to measure social network structures, and influence, relative to the secondary analysis in the previous chapter. Peer influences can be measured in many ways, but I focus on source credibility and tie strength. Tie strength represents the closeness between individuals regarding while source credibility represents the perceived expertise on the subject. I measure source credibility by collecting data on how subjects regard the apparent expertise of others, while I measure tie strength by asking how well each subject knows the others.

By populating social weight matrices with each measure, I then estimate the importance of each in shaping endogenous peer effects. The relative effectiveness of each of these mechanisms is critical to the conduct of modern marketing mechanisms. While traditional word-of-mouth (WOM) marketing relies on suasion from proximal consumers such as family and friends, internet-based marketing tools rely more on the credibility of third-party opinions voiced by “expert” sources.

I find that perceived source credibility is more important in influencing preferences than the closeness of social relationships. This finding is consistent with Freeman (1957) and Pornpitakpan (2004) who maintain that sources having more credibility are more influential than are those with less credibility. Moreover, in contrast to the findings of Richards, Hamilton and Allender (2014), people who are close in social space are not likely to have a significant influence on each other because they already share similar information. My findings, therefore, echo the “strength of weak ties” effect documented by Granovetter (1973) in that “weak ties” convey more influential product information that “strong ties”.

In particular, because subjects who are perceived as credible are not necessarily close friends, marketers should target individuals who have perceived credibility on the product of interest when promoting innovative products. Traditional marketing focuses on word-of- mouth between acquaintances. However, my results show that consumers do not necessarily take advice from people they are close to. Whether explicitly or not, marketers leverage social media to this effect already, because viral marketing practices such as Yelp and TripAdivsor gather recommendations from strangers who have experience with a particular product or service. Strangers, by definition, do not have social ties with the decision maker, so are perceived as credible sources. More fundamentally, my findings support the underlying logic of the weak-ties model in a modern marketing context: Consumers are sufficiently rational in that they implicitly understand that social proximity does not necessarily imply similar tastes, so that expert opinion can do more to reveal true congruence between product attributes and individual tastes than the opinions of "friends."

My findings are important to marketing practice because marketers are always seeking ways to introduce new products to the market (Ferguson 2008, Bruyn and Lilien 2008). In this regard, I highlight an essential point of the social learning literature, namely that how relationships are defined is essential to understanding the nature and power of social influence within a network.

The rest of this dissertation is arranged in the following fashion. In the first essay (chapter 2) I examine the optimal pricing of agricultural innovation. The second essay (chapter 3) studies the strategic behavior in variety adoption. Next, I investigate consumers' choices in a peer network in the third essay (chapter 4). I reserve my concluding remarks, and offer more general implications of my findings, for a final chapter.

CHAPTER 2. ESSAY 1: FEES VERSUS ROYALTIES IN AGRICULTRAL PATENTS

2.1 Introduction

Our formal understanding of the optimal mechanism for patent licensing has changed considerably in recent years. Kamien and Tauman (1984, 1986), Katz and Shapiro (1985), and Kamien, Oren, and Tauman (1992) find that licensing via a royalty system generates less revenue for an external innovator than if a fixed fee or auction were used. However, the empirical research in non-agricultural industries tends to find that royalties, or combinations of fees and royalties, are far more common (Sen and Tauman, 2007). The challenge facing researchers then became reconciling this stylized fact with economic theory. By including more realistic institutional attributes of industry such as product differentiation (Motta, 1993; Faulí-Oller and Sandonís, 2002), asymmetric information (Gallini and Wright, 1990; Sen, 2005), risk aversion (Bosquet et al. 1998), moral hazard (Choi, 2001), incumbency (Shapiro, 1985; Kamien and Tauman, 2002; Wang, 2002; Sen and Tauman, 2007) or strategic delegation (Saracho, 2002) researchers were able to explain observed licensing strategies. Faulí-Oller and Sandonís (2002), for example, show that regardless of the type of competition, the optimal contract always includes a positive royalty when products are differentiated. Our challenge, therefore, is to explain why fees tend to dominate in the context of horticultural innovations.

Most of the theoretical literature on licensing patented research concerns cost-reducing innovations. In agriculture in general, and in the fruit and vegetable sector more specifically, however, a significant number of innovations seek to improve eating quality, a demand-side innovation. Unlike a cost-reducing innovation, a quality-improving innovation directly affects consumers' preferences and their willingness to purchase a product. Among studies that consider demand-side or product innovations, Bosquet et al (1998) find that a combination of fees and royalties is optimal if demand for the new product is uncertain. In their model, fees and royalties are a means by which a risk neutral innovator can provide insurance--and be compensated for it--to a risk-averse licensee. Sen (2005) generates a similar combination of tools under asymmetric information. If the licensee has private information regarding its cost of producing the new product, then the licensor will benefit from using a combination of fees and royalties. Li and Wang (2010) consider a Cournot duopoly scenario in which the external innovator sells a quality improving innovation and find that exclusive licensing is preferred under fixed fees while non-exclusive licensing is preferred under royalties and two-part tariffs. They, however, do not consider R&D cost incurred by developing the innovation.

I consider the strategic rationale for pricing a demand-side innovation into a downstream Bertrand duopoly market. Adopting a discrete choice-modeling framework (Anderson, de Palma, and Thisse, 1992) to study the behavior of oligopolies under product differentiation, the innovator licenses its output using a fixed fee, royalty, or combination of the two. When the market is covered (all consumers buy), they find that both firms purchase the innovation by paying a positive royalty and no fixed fee. If the value of the outside option is relatively high, then both firms will still license the innovation but pay a combination of fee and royalty. Although they show that quality-enhancing innovations are licensed using a contract that includes both fees and royalties, they do not treat the degree of innovation as a continuous variable. Therefore, it is not clear whether their result holds regardless of whether innovations are both minor and significant. This paper derives threshold values (for the degree of innovation) that define whether fees, royalties, or both are optimal.

This chapter offers a theoretical model of optimal licensing schemes for quality-improving agricultural innovations. I consider an oligopolistic market where two downstream firms compete in price and the upstream innovator holds a quality-improving technology that may create differentiation between the products.[2] Since I am interested in university-based research specific to (but not limited to) the horticultural industry, the innovator is an outsider by default. I consider both exclusive and non-exclusive licensing. Under exclusive licensing, only one downstream firm gets the innovation, and in non-exclusive licensing, more than one firm is allowed to produce and sell the new product. This framework provides a realistic yet tractable description of the market for demand-side agricultural innovations.

I find that the innovator maximizes licensing revenue under a non-exclusive, fixed-fee regime. In general, the results show that non-exclusive licensing performs better than exclusive licensing under both fixed fees and royalties and that a two-part tariff scheme will not be used because neither downstream firm can improve upon their pre-license profit level. With a fixed fee, the innovator is able to extract the licensing firms' increased profits but is not able to control industry output. Licensing through a royalty, the innovator is able to manipulate the cost structure of the licensing firms, which provides a measure of control over the final output. Two-part tariffs have the potential to generate the most revenue, but I find that licenses will never be obtained this way. When

[2] Price competition is not necessary for firms that sell differentiated products, but given that most produce is sold through retail stores, price competition is a more natural choice.

the innovator has control over the market, it is in her best interest to intensify competition between the downstream firms by licensing to one firm and then extracting rents generated by the market power conferred on the higher-quality producer. When the innovator does not have control over the final output, it is in her best interest to license to both firms and collect as much additional profit as possible from the innovation through a fixed fee. Further, licensing through either a fixed fee or a two-part tariff moderates competition between two downstream firms and results in a market that produces only high-quality products.

2.2 Model

I consider a final market with two firms and two differentiated products: High quality products and low quality products, where high quality products are produced with the innovation and low quality products are produced through the existing technology. Under exclusive licensing, I assume each firm produces only one type of product. Competition in the final market results from the firms selling differentiated products. The innovation is patent-protected. Three types of licensing contract are considered in this chapter: (i) a fixed fee based license, where the licensee pays F to the patent holder regardless of the quantity he will sell in the final market, (ii) a royalty-based license where the licensee pays r to the patent holder for each unit he will sell, and (iii) a combination of both payment schemes where the licensee pays both a fixed up-front fee F and a per unit royalty r for the quantity sold. I assume an oligopoly that consists of two firms, each producing a differentiated good.

On the consumer side, I assume a continuum of consumers of the same type with a utility function separable and linear in each good, which facilitates partial equilibrium analysis. That is, a representative consumer maximizes a quadratic, strictly concave utility function, which gives rise to a linear demand structure. Consumers are willing to pay more for higher quality products, where the maximum willingness to pay is given by $c(s_i)$. Differentiation comes from two sources: the degree of substitutability, b; and the quality, s_i. Following Sign and Vives (1984), the representative consumer maximizes:

$$U(q_1, q_2) - \sum_{i=1}^{2} p_i q_i \tag{2.1}$$

U is assumed to be quadratic and strictly concave where:

$$U(q_1, q_2) = c(s_1)q_1 + c(s_2)q_2 - \frac{1}{2}(q_1^2 + 2\frac{b}{s_i}q_1 q_2 + q_2^2) \tag{2.2}$$

Therefore, inverse demand for each product on the downstream market is

$$p_i = c(s_i) - q_i - (b/s_i)q_j \qquad (2.3),$$

where p_i is the price set by firm i: i = 1, 2; and $q_i$represents the quantity sold by firm i. Without loss of generality I assume that firm 1 produces low quality products and firm 2 produces high quality products under exclusive licensing and that both firms produce high quality products under non-exclusive licensing.

The variable s_i measures quality, with s_1 indicating low quality and s_2 indicating high quality. The highest propensity to pay for quality s_i is denoted by $c_i = c(s_i)$. Following Li and Wang (2010), I first normalize s_1 to be 1 and then assume a relationship between low quality and high quality where $s_1 = \lambda s_2(\ \lambda \in (0,1))$, where λ captures the degree of product innovation: A larger λ implies a smaller quality improvement and a smaller λ indicates a greater quality improvement. I assume that λ is exogenous, which reflects the fact that TTOs are charged with marketing innovations that are presented to them from their faculty/innovators.

The degree of substitutability between the products is indicated by b (b ∈ (0, 1)). When b = 1 and λ = 1 the two products are perfect substitutes. Including both is necessary to isolate the quality-enhancing nature of innovations. Namely, the parameter b captures the fact that the products are horizontally differentiated, or that there is at least part of the market that would prefer each product even if the prices were the same. On the other hand, λ introduces a vertical component in that the willingness-to-pay for high-quality goods rises in $1/\lambda$ for the entire market. In the absence of the b parameter, the Singh and Vives (1984) model has no way of separating an innovation that is truly better, from one that is merely different. By introducing both parameters, I separate the two effects, and base our licensing model on a more general demand framework. Further, I assume a quadratic structure for the highest propensity to pay ($c(s_i) = s_i^2$) in order to ensure an interior solution (Sen and Tauman 2007).

Under non-exclusive licensing, both firms face similar demand functions and produce either low or high products exclusively in the final market. Under exclusive licensing, the demand functions facing low and high quality firms are, respectively:

$$q_1 = \frac{1 - \mathrm{p}_1 + bp_2 - \frac{b}{\lambda^2}}{1 - b^2\lambda} \tag{2.4}$$

$$q_2 = \frac{-p_2 + \frac{1}{\lambda^2} - b\lambda + bp_1\lambda}{1 - b^2\lambda} \quad (2.5),$$

where the own price response for low quality products is $\frac{-1}{1-b^2\lambda}$ and the cross price response for low quality products is $\frac{-b}{1-b^2\lambda}$. The own price response for high quality products is $\frac{-1}{1-b^2\lambda}$ and the cross price response for high quality products is $\frac{b\lambda}{1-b^2\square}$. Both intrinsic characteristics b and quality λ play roles in differentiating low and high products. I focus on the impact of quality on price differentiation.

2.3 Propositions Regarding the Innovator's Profit

In this section, I study optimal licensing strategies for the innovator.[3] To do so, I first consider the profitability of each downstream firm and analyze their incentive to license the innovation. When the patent is licensed exclusively to one firm we refer to the licensee as firm 2. Throughout this paper, $p^{ij}{}_k$, $q^{ij}{}_k$ and $\pi_k{}^{ij}$ denote firm k's price, quantity, and profit by means of contract i, where k = 1 is firm 1 (low quality), k = 2 is firm 2 (high quality), and k = 3 is the innovator; $j = E$ indicates an exclusive contract; $j = N$ is a non-exclusive contract; and $i = NL,FE,FN,RE,RN,TE,TN,$ which represent, respectively, no licensing, exclusive fixed-fee licensing, non-exclusive fixed-fee licensing, exclusive royalty licensing, non-exclusive royalty licensing, exclusive two-part tariff licensing, and non-exclusive two-part tariff licensing. For example, p^{FN} is the market price when the innovation is licensed to both firms through fixed fee. The innovator's profit is the sum of

[3] Optimality is defined, as in the literature, as the difference between license revenues and the cost of innovation. Although our analysis concerns university research activities, and universities are expected to conduct basic research in the public interest, our profit-maximization assumption reflects the observed activities of university TTO offices. Namely, as Bulut and Moschini (2009) note, "Quite clearly, when it comes to patenting and licensing, universities are likely to behave based on their self-interest rather than the public interest" (p. 124). Resolving the debate as to whether universities should maximize the returns to their research investment is left for either political or legal discussion.

any royalty or fee less the cost of innovation. I assume the cost of innovation is convex in the extent of the quality improvement and assumes the same form as the propensity to pay, or $c(s_i) = {s_i}^2$

The licensing game consists of three stages. In the first stage, the innovator simultaneously offers either non-exclusive contracts, or an exclusive contract, consisting of royalties, fees, or a combination of the two. In the second stage, the downstream firms either accept or reject the license contracts. In the third stage, the firms compete in the downstream market. I first consider the case where no license is purchased in order to calculate the benchmark profit for both firms under an exclusive licensing scenario. When licensing is non-exclusive, I first solve for the optimal solution to the sub-game played between downstream firms in order to establish the benchmark profit. The benchmark profit becomes the profit of the licensing firm under exclusive licensing. Then I consider each of the other licensing strategies in the following order: Fixed fee licensing, royalty licensing, and two-part tariff licensing. I compare profits under each licensing strategy with the benchmark profit, and suggest the optimal licensing strategy for the patent holder. I also study the sub games among the downstream firms and compare their profit under both price competition and quantity competition and suggest optimal strategies played by the downstream firms.

No Licensing

I establish benchmark profits where no innovation is introduced. In this case, firms produce only low-quality products. Following Motta (1993), I assume constant marginal costs and normalize them to be 0. Therefore, the duopoly profits when no license is purchased are

$$\pi_1^{NL} = max p_1 q_1 = p_1[\frac{1}{1-b^2}(-p_1 + bp_2 - b + 1)] \tag{2.6}$$

$$\pi_2^{NL} = max p_2 q_2 = p_2[\frac{1}{1-b^2}(-p_2 + bp_1 - b + 1)] \tag{2.7}$$

Solving the first order conditions of this problem results in the optimal prices:

$$p^{NL} = p_1^{NL} = p_2^{NL} = -\frac{1-b}{-2+b} \tag{2.8}$$

In this expression, $p^{NL} > 0$ because $1 - b$ must be positive and $-2 + b$ must be negative as b is between 0 and 1. Price competition under no licensing results in positive market prices. Profit is symmetric and depends solely on the degree of substitutability between the products. The expression for the profit earned by both firms becomes:

$$\pi^{NL} = \pi_1^{NL} = \pi_2^{NL} = \frac{1-b}{(-2+b)^2(1+b)} \tag{2.9}$$

This is also positive because b is between 0 and 1. Both firms make positive profits under the no licensing scenario and when they both produce the low-quality products. I refer to such a profit as the benchmark profit. When the innovation is introduced into the market through licensing, the demands for high- and low-quality products will change and so will firms' profits. Firms compare their potential profits with the benchmark profit and decide whether it is in their best interests to license the innovation.

Fixed-Fee Licensing

When licensing using fixed fees, the patent holder extracts the entire profit due to the innovation by setting the fixed fee equal to the difference between the licensee's profit with the innovation and the benchmark profit. If the fee were any larger, the licensees would be better off without the patent as the new profit will be smaller than the benchmark profit. If the fixed fee is smaller than the incremental profit, the innovator will not extract all the profit and can always benefit more by increasing the fixed fee until it is exactly equal to the difference.

PROPOSITION 1. *Under exclusive fixed fee licensing, the innovator makes a positive profit when λ is smaller than 0.40.*

Proof. Under exclusive fixed-fee licensing, only one firm purchases the patent. I assume that firm 2 purchases the innovation and produces high-quality products. The profit is given by

$$\pi_2^{FE} = p_2\left[\frac{1}{1-b^2\lambda}\left(-b\lambda p_1 - p_2 + \frac{1}{\lambda^2} - b\lambda\right)\right] - F^{FE} \tag{2.10}$$

Since firm 1 doesn't purchase the patent, it does not yield any revenue directly to the innovator, but its optimal price conditions the profit from firm 2, the firm purchasing the patent. In this case, firm 1 produces low-quality products with a profit of:

$$\pi_1^{FE} = p_1\left[\frac{1}{1-b^2\lambda}\left(-p_1 + bp_2 - \frac{b}{\lambda^2} + 1\right)\right] \tag{2.11}$$

Solving for the optimal fee and subtracting the benchmark profit leaves a fixed fee of

$$F^{FE} = \frac{1-b}{(-2+b)^2(1+b)} - \frac{(-2+b^2\lambda+b\lambda^3)^2}{\lambda^4(-4+b^2\lambda)^2(-1+b^2\lambda)} \tag{2.12}$$

Therefore profit for the innovator becomes:

$$\pi_3^{FE} = F^{FE} - c_2 = \frac{1-b}{(-2+b)^2(1+b)} - \frac{(-2+b^2\lambda+b\lambda^3)^2}{\lambda^4(-4+b^2\lambda)^2(-1+b^2\lambda)} - \frac{1}{\lambda^2} > 0 \tag{2.13),}$$

where $c_2 = c(s_2) = 1/\lambda^2$ is the cost of innovation. This expression is only positive when $\lambda < 0.40$ (at $b = 0.50$), so the innovator makes a positive profit under exclusive fixed-fee licensing when the innovation is substantial (see figure 2.1) at a moderate level of substitutability. By using the fixed-fee strategy, the innovator is able to extract all the profit above benchmark profit, leaving the profit of licensee exactly equal to the benchmark profit.

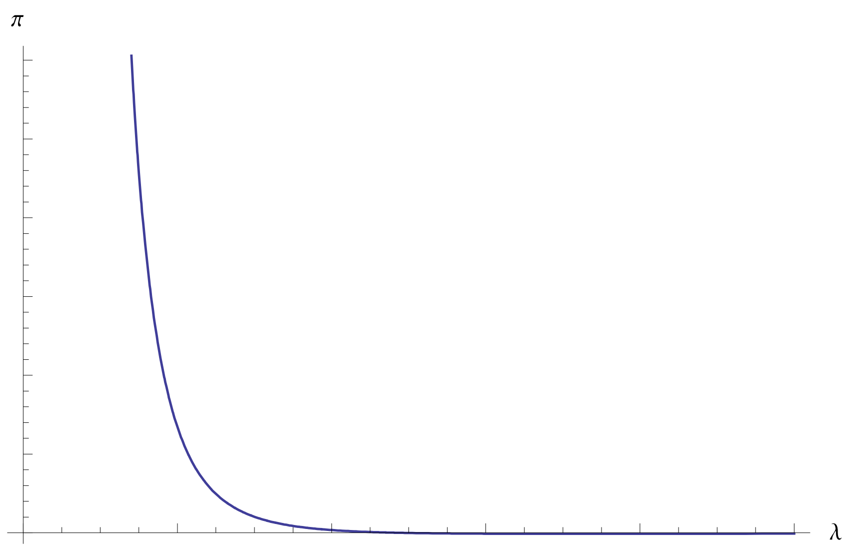

Figure 2.1 Innovator Profit under Exclusive Fixed-Fee Licensing

PROPOSITION 2. *Under a non-exclusive fixed-fee strategy, the innovator makes a positive profit when* $\lambda < 0.50$.

Proof. Under non-exclusive licensing, both firms purchase the patent and produce high-quality products. In this case, however, the benchmark profit for both firms is the profit of the low-quality firm under exclusive licensing, which is the same as the profit under no licensing. Since the innovator licenses through a fixed fee to both, she is able to extract the extra profit of both firms and leave them with benchmark profits. The profits for both firms are written as

$$\pi_1^{FN} = p_1^{FN}\left[\frac{1}{1-b^2\lambda^2}\left(-p_1 + b\lambda p_2 + \frac{1}{\lambda^2} - \frac{b}{\lambda}\right)\right] - F^{FN} \tag{2.14}$$

for firm 1 and:

$$\pi_2^{FN} = p_2^{FN}\left[\frac{1}{1-b^2\lambda^2}\left(-p_2 + b\lambda p_1 + \frac{1}{\lambda^2} - \frac{b}{\lambda}\right)\right] - F^{FN} \tag{2.15}$$

for firm 2. Solving both maximization problems results in a fixed fee of:

$$F^{FN} = \frac{4 - 4\lambda^4 + 4b\lambda(\lambda^3 - 1) + 3b^2(\lambda^6 - 1) - b^3(\lambda^7 + 3\lambda^6 - 3\lambda - 1)}{(b-2)^2(b+1)\lambda^4(b\lambda-2)^2(b\lambda+1)} \tag{2.16}$$

So the level of profit for the innovator becomes

$$\pi_3^{FN} = 2F^{FN} - c_2 = \frac{2(4 - 4\lambda^4 + 4b\lambda(\lambda^3 - 1) + 3b^2(\lambda^6 - 1) - b^3(\lambda^7 + 3\lambda^6 - 3\lambda - 1))}{(b-2)^2(b+1)\lambda^4(b\lambda-2)^2(b\lambda+1)} - \frac{1}{\lambda^2} \tag{2.17}$$

The innovator licenses through a fixed fee to both firms, so she is able to extract the extra profit from each, which leaves them with profits equal to the benchmark. Because innovation is costly, however, I again observe a threshold level of quality above which innovation will not make sense from the upstream firm's perspective. Again, fixing the level of $b = 0.5$ for comparison purposes, I calculate profit under a range of λ values as shown in figure 2.2.

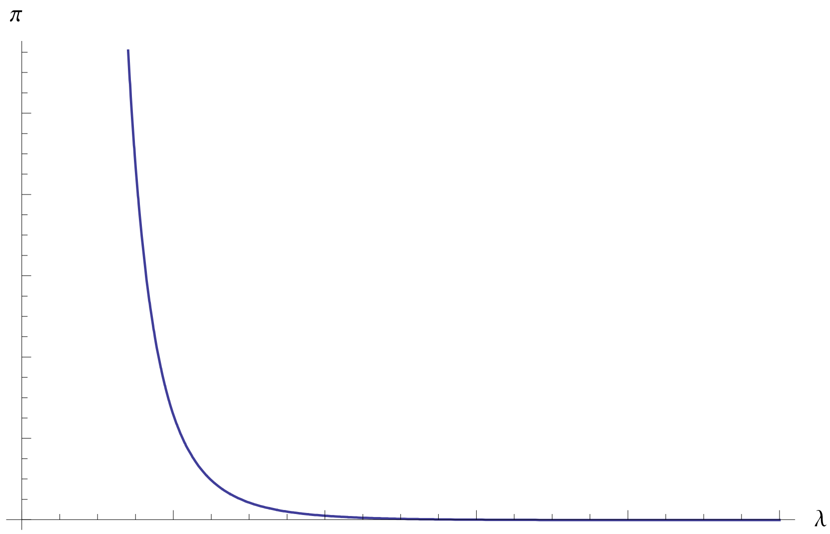

Figure 2.2: Innovator Profit under Non-Exclusive Fixed-Fee Licensing

Under non-exclusive fixed-fee licensing, both firms produce high-quality products. After extracting the increased profits from both firms and compensating for the cost incurred by investing in the innovation, the innovator only makes a positive profit when $\lambda < 0.5$. Since there are incentives to license the patent under both exclusive and non-exclusive licensing, a comparison of the innovator's profits under both scenarios will yield a better understanding of the optimal licensing under a fixed-fee strategy.

PROPOSITION 3. *Under a fixed-fee strategy, the patent-holding firm prefers non-exclusive licensing.*

Proof. To understand which licensing strategy is better under a fixed fee, I take the difference between profits under non-exclusive and exclusive licensing arrangements:

$$\pi_3^{FN} - \pi_3^{FE} = \frac{-1+b}{(-2+b)^2(1+b)} + \frac{(-2+b^2\lambda+b\lambda^3)^2}{\lambda^4(-4+b^2\lambda)^2(-1+b^2\lambda)} - \frac{2\left(4-4\lambda^4+4b\lambda(\lambda^3-1)+3b^2(\lambda^6-1)\right)}{(b-2)^2(b+1)\lambda^4(b\lambda-2)^2(b\lambda+1)} + \frac{b^4\lambda(\lambda^6-1)-b^3(\lambda^7+3\lambda^6-3\lambda-1)}{(b-2)^2(b+1)\lambda^4(b\lambda-2)^2(b\lambda+1)} > 0 \tag{2.18}$$

This result indicates that non-exclusive licensing yields larger profits for the innovator. Under fixed-fee licensing the patent-holding firm is willing to license her patent to both firms instead of just one. The intuition behind this proposition is straightforward. The innovator is able to extract all of the extra profit from the innovation by charging a fixed fee to both firms, leaving the profits of the licensees exactly equal to the benchmark profit. This finding is contrary to Li and Wang (2010), who find that the patent holder is willing to sell its patent to a single firm under a fixed-fee contract. Li and Wang (2010) consider a Cournot duopoly framework in which firms compete in quantities. In their model, a non-exclusive licensing strategy was not preferred because licensing to both firms generates the same quality improvement without affecting competition. I consider instead a Bertrand duopoly framework in which firms compete in prices. When firms sell differentiated products and compete in prices, the innovation generates higher demand at a higher price level. The innovator is better off licensing her patent to both firms, thus clearing low-quality products out of the market.

Royalty Licensing

Royalties are different from fixed fees in that the innovator cannot extract all of the downstream profit through a royalty scheme but can better preserve industry profit by changing downstream firms' output. Because the innovator's profit is positively related to output, it can generate greater license revenue by incentivizing higher industry output. I first consider exclusive licensing then non-exclusive licensing.

PROPOSITION 4. *Under exclusive royalty licensing, the innovator makes a positive profit when the level of innovation is high (*$\lambda \leq 0.35$*).*

Proof. Under exclusive royalty licensing, the innovator sells her patent to only one firm. Without loss of generality, we assume that firm 2 purchases the patent and produces high-quality products. The royalty becomes part of the marginal cost, denoted by r. The profit function for firm 2 is revenue after accounting for the royalty payment and is written as

$$\pi_2^{RE} = (p_2^{RE} - r^{RE})q_2^{RE} \qquad (2.19)$$
$$= (p_2 - r)[\frac{1}{1-b^2\lambda}\left(-p_2 + b\lambda p_1 + \frac{1}{\lambda^2} - b\lambda\right)]$$

Firm 1 then produces low-quality products. Because this firm does not purchase the patent, its profit is irrelevant to the income earned by the innovator, but its optimal price conditions the profit earned by firm 2. Firm 1's profit derives from selling only low-quality products, so its optimal choice of price is found as the solution to

$$\pi_1^{RE} = p_1^{RE}q_2^{RE} = p_1[\frac{1}{1-b^2\lambda}\left(-p_1 + bp_2 - \frac{b}{\lambda^2} +\right)] \qquad (2.20)$$

The innovator earns a per unit royalty for every unit sold by firm 2, so by deducting the cost of innovation from the revenue earned from firm 2, I obtain the innovator's profit as

$$\pi_3^{RE} = r^{RE}q_2^{RE} - \frac{1}{\lambda^2} \qquad (2.21)$$

Solving for the optimal royalty rate gives

$$r^{RE} = \frac{-2 + b^2\lambda + b\lambda^3}{4\lambda^2(-2 + b^2\lambda)} > 0 \qquad (2.22)$$

Substituting this expression back into the profit functions yields[4]

[4] Detailed derivations are provided in the appendix A.

$$\pi_3^{RE} = \frac{-4b^6\lambda^5+b^4\lambda^2(28\lambda^2-1)-2b^3\lambda^4-b^2\lambda(\lambda^5+56\lambda^2-4)+4b\lambda^3+32\lambda^2-4}{4\lambda^2(-2+b^2\lambda)(b^2\lambda-4)(b^2\lambda-1)} \tag{2.23}$$

To sign this expression, I again fix b = 0.5 and calculate the relationship between λ and innovator profit (see figure 2.3). The innovator thus earns a positive profit when degree of innovation is relatively high ($0 < \lambda \leq 0.35$) and a negative profit when degree of innovation is relatively low ($0.35 < \lambda < 1$).

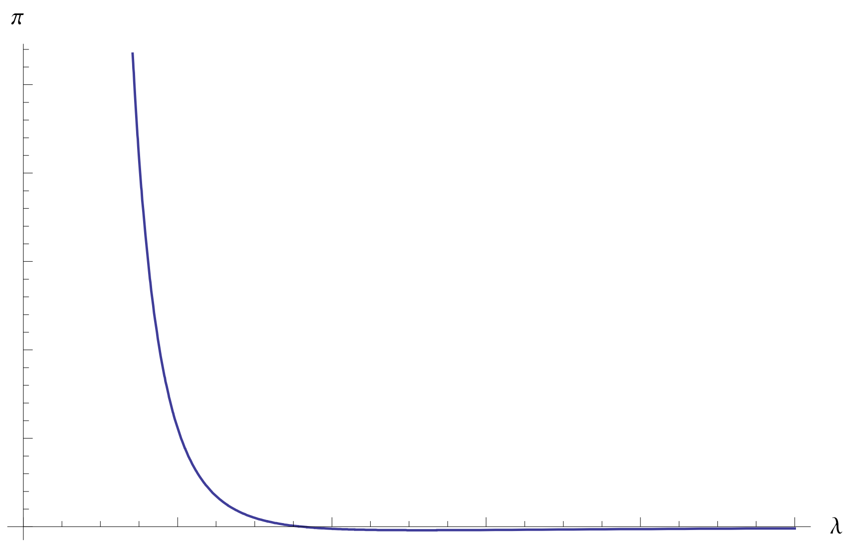

Figure 2.3: Innovator Profit under Exclusive Royalty Licensing

Our results with respect to royalty contracts are intuitive because only a significant innovation should generate positive returns to the innovator. To see this, consider that the introduction of royalties has two effects. First, the royalty becomes part of marginal cost, which increases the price of the high-quality product. Second, the royalty can influence output in the downstream market. Under exclusive licensing, the innovator's profit is closely related to the profit of the high-quality firm. The innovator wants to set a royalty that is low enough to induce higher output yet not too low, as royalties are earned on a per unit basis. When the innovation is sufficiently large ($0 < \lambda \leq 0.35$), the innovator benefits due to higher output, but when the innovation is relatively small (0.35 <λ < 1), the innovator earns less because the loss of demand dominates the higher profit incurred by greater differentiation.

PROPOSITION 5. *Under non-exclusive royalty licensing, the innovator makes a positive profit when the level of innovation is high (*λ < 0.50).

Proof. Under non-exclusive licensing, the innovator sells the patent to both firms and controls the entire output through royalty licensing. The profit functions are given by:

$$\pi_2^{RN} = (p_2^{RE} - r^{RE})q_2^{RE} \tag{2.24}$$
$$= (p_2 - r)[\frac{1}{1 - b^2\lambda^2}\left(-p_2 + b\square p_1 + \frac{1}{\lambda^2} - \frac{b}{\lambda}\right)$$

for firm 2 and:

$$\pi_1^{RN} = (p_1^{RE} - r^{RE})q_1^{RE} = (p_1 - r)[\frac{1}{1 - b^2\lambda^2}\left(-p_1 + b\lambda p_2 + \frac{1}{\lambda^2} - \frac{b}{\lambda}\right) \tag{2.25}$$

for firm 1. Profit to the innovator is given by:

$$\pi_3^{RN} = r^{RN}(q_1^{RN} + q_2^{RN}) - \frac{1}{\lambda^2} \tag{2.26}$$

Again, solving for the optimal royalty gives

$$r^{RN} = \frac{1}{2\lambda^2} > 0 \tag{2.27}$$

Substituting the result back into the innovator's profit function and solving leave

$$\pi_3^{RN} = \frac{1 - 4\lambda^2 - 2b\lambda^3 + 2b^2\lambda^4}{4\lambda^4 + 2b\lambda^5 - 2b^2\lambda^6} \tag{2.28}$$

Signing this expression is again difficult analytically, so I calculate innovator profit at $b = 0.5$ and show how profit varies with λ (see figure 2.4).

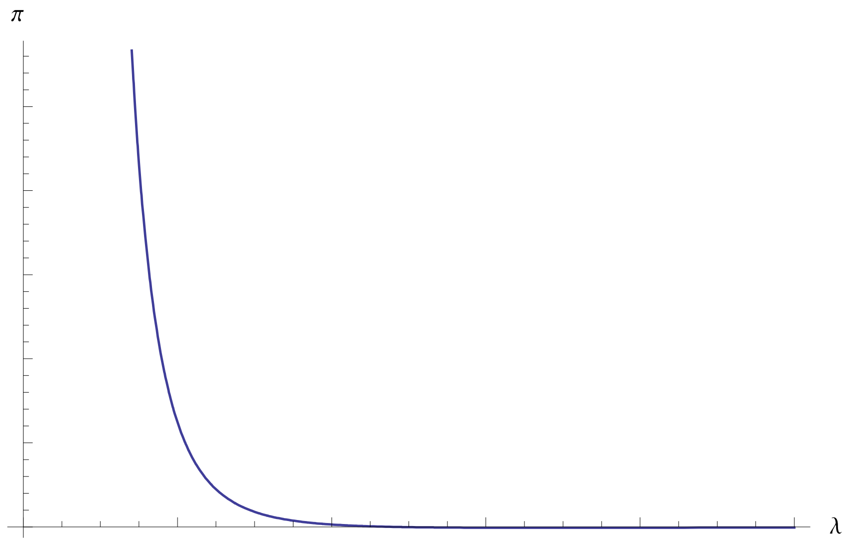

Figure 2.4: Innovator Profit under Non-Exclusive Royalty Licensing

Figure 4 shows that the innovator has an incentive to license when $\lambda < 0.50$. Even though both firms produce the high-quality product and face similar demand functions, their profit differs from the benchmark profit (unlike in the case of non-exclusive fixed-fee licensing) because the royalty alters the structure of demand as it increases marginal cost. The overall profit from the high-quality market under non-exclusive licensing depends not only on inherent product differentiation (b) but also on the differentiation brought by innovation (λ). When the magnitude of the innovation is larger, the innovator's profit rises, and when the magnitude of innovation is smaller, the innovator's profit falls. In the next proposition, I compare the profits earned under both strategies to get a better understanding of how royalty alters the innovator's profit.

PROPOSITION 6. *Under a royalty contract, the patent holder favors non-exclusive licensing.*

Proof. The difference in profits between exclusive licensing and non-exclusive licensing is[5]

$$\pi_3^{RN} - \pi_3^{RE} > 0 \qquad (2.29),$$

which is greater than zero. This result indicates that non-exclusive royalty licensing yields greater profit for the innovator relative to exclusive licensing. Compared with the situation of licensing to one firm, licensing to both firms generates higher aggregate output, which leads to higher licensing profit. As a consequence, the patent holder is willing to transfer its technology to both. This finding differs qualitatively from the outcome expected by Li and Wang (2010) under quantity competition, as they favor exclusive licensing. In their case, quality enhancement creates asymmetric demands between the two firms, softening market competition and generating higher incremental profit for the high-quality firm. With royalty licensing, the innovator's profit is directly related to that earned by the high-quality firm. Price competition, on the other hand, favors increasing output from both firms.

Two-Part Tariff Licensing

When comparing a fixed fee with a royalty, I see that with a fixed fee the innovator is able to extract a lump sum of profit above the benchmark profit without changing the nature of competition between the firms. By licensing to both firms with a fixed fee, the innovator can set the licensees' profits back to the benchmark level and they will still have an incentive to purchase the license. With a royalty, however, the structure of competition is changed because higher royalties raise marginal cost.

[5] Detailed derivations are provided in the appendix A.

Therefore, even when the innovator licenses to both firms, their profits differ from the benchmark level. In the following section, I show how a combination of both fixed fees and royalties affect the innovator's profit.

Licensing by a two-part tariff is more complicated because the patent holder must trade off two effects: On one hand, the patent holder has an incentive to lower the royalty in order to moderate competition between the downstream firms and preserve industry profit, then extract it with fixed fees. On the other hand, the patent holder has an incentive to keep the royalty higher in order to extract as much profit as possible from the licensees. The more profound the innovation, the lower the net profit of the licensee and the better off the licensor.

PROPOSITION 7. *Under exclusive two-part tariff licensing, firm 1 makes more profit than the benchmark level, while firm 2 makes less profit than the benchmark, so licensing will not occur.*

Proof. By using a two-part tariff, the innovator sets a royalty to control output in the final market and a fixed fee to extract any excess profits. Under exclusive licensing, I assume the innovator sells her patent to only firm 2, therefore firm 2 produces high-quality products and firm 1 produces low-quality products. Recall from the nature of the game that the innovator's optimal decision is conditional on the solution to the subgame played among the downstream firms. Profitable licensing depends on the willingness of at least one firm to purchase the license. To see why neither will, consider the profit earned by firm 2[6]:

[6] Detailed calculations are in Appendix A.

$$\pi_2^{TE} = (p_2^{TE} - r^{TE})q_2^{TE} < \pi^{NL} \tag{2.30}$$

and

$$\pi_1^{TE} = p_1^{TE} q_1^{TE} > \pi^{NL} \tag{2.31}$$

by firm 1. Solving for the optimal royalty gives a value of

$$r^{TE} = -\frac{b^2\lambda + b\lambda^3 - 2}{\lambda^2(b^2\lambda - 6)(b^2\lambda - 2)} > 0 \tag{2.32}$$

Combining royalty and fee, the maximum innovator profit becomes[7]

$$\pi_3^{TE} = r_2^{TE} q_2^{TE} + F^{TE} \tag{2.33}$$

which we find to be positive when $\lambda < 0.4$ for a fixed value of $b = 0.5$ (see figure 2.5). Even though the innovator makes a positive profit when $\lambda < 0.4$, the profit will not be realized because firm 2 makes less profit than the benchmark and, therefore, will not purchase the patent.

[7] Detailed calculations are in Appendix A.

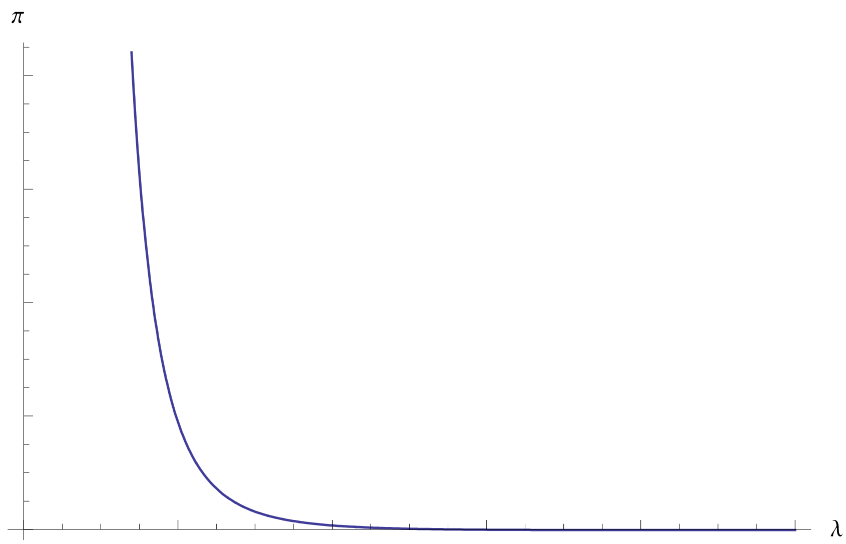

Figure 2.5: Innovator Profit under Exclusive Two-Part Tariff Licensing

PROPOSITION 8. *Under non-exclusive two-part tariff licensing, both firm 1 and firm 2 make less profit than the benchmark level, so licensing will not occur.*

Proof. Under non-exclusive licensing, the innovator sells the patent to both firms. Both firms produce high-quality products, and low-quality products are cleared out of the market. The licensing firms benefit from producing high-quality products but are required to pay a per unit royalty and an up-front fixed fee. The profit functions for the two firms are now written as

$$\pi_j^{TN} = \left(p_j^{TN} - r^{TN}\right)q_j^{TN} - F^{TN} \tag{2.34}$$

for firm j = 1,2. The fixed fee is set to equal the difference between new profit and the benchmark profit. Since both firms buy licenses and the returns are symmetrical, the fixed fee is the same for both firms:

$$F^{TN} = \pi_2^{TN} - \pi^{NL} \quad (2.35)$$

Solving for the optimal royalty gives:

$$r^{TN} = \frac{1}{6\lambda^2 - 2b\lambda^3} > 0 \quad (2.36)$$

Substituting the optimal royalty and fee expressions back into the symmetric profit functions, I find[8]

$$\pi_1^{TN} = \pi_2^{TN} < \pi^{NL} \quad (2.37),$$

therefore neither firm is willing to purchase a license.

Hypothetically, the innovator thus earns equal up-front fixed fees from both firms and per unit royalty payments for every unit produced. Profit for the innovator is given by the solution to[9]

$$\pi_3^{TN} = 2F_{TN} + r^{TN}(q_1^{TN} + q_2^{TN}) - \frac{1}{\lambda^2} \quad (2.38),$$

which would be positive if the downstream firms choose to purchase the patent in the second stage of the game. As in the case of non-exclusive royalty licensing, both firms produce high-quality products and face the same demand. Their profits differ from the benchmark level because the royalty alters the structure of demand. Total profit in the high-quality market under non-exclusive licensing depends not only on the fact that the

[8] Detailed calculations are in Appendix A.

[9] Detailed calculations are provided in Appendix A.

products are differentiated from each other (*b*) but also the vertical differentiation due to the innovation (λ). When λ is smaller (larger innovation), innovator profit is higher because the margins earned downstream by the purchasing firms are larger. Since both profits under exclusive and non-exclusive licensing are positive, I take the difference between the two to determine the optimal licensing strategy.

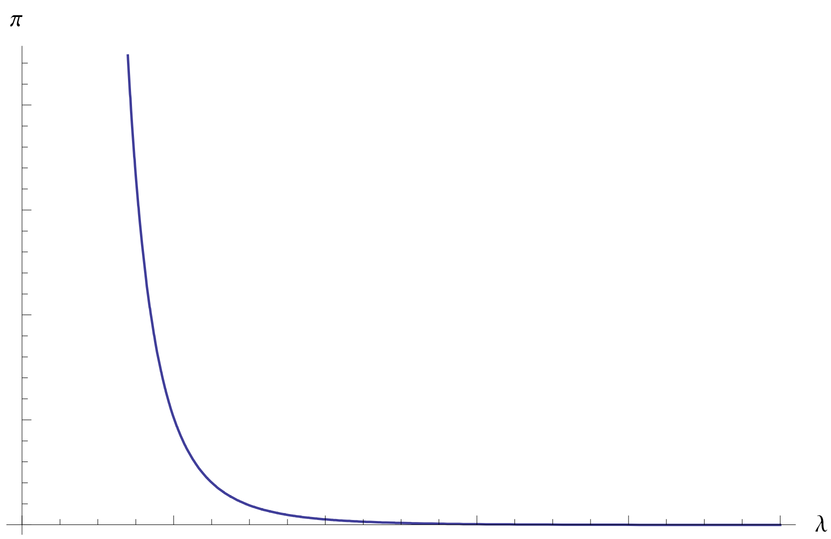

Figure 2.6: Innovator Profit under Non-Exclusive Two-Part Tariff Licensing

Comparing innovator profit between exclusive and non-exclusive contracting with a two-part tariff is therefore meaningless because licensing will not occur in either case.

2.4 Conclusion

Downstream firms that license patents to new products have an incentive to purchase a license as long as the new profit brought by the innovation exceeds the benchmark profit, or the profit implied by the equilibrium to the second-stage of the game played downstream. After choosing whether to use a royalty, fee, or a combination of the two, the innovator faces the option of either selling to one firm or to both firms. Whether one or two firms purchase the license changes the fundamental structure of the market because under non-exclusive licensing both firms produce high-quality products and the low-quality products will be forced from the market. Because the innovator's decision is driven by the willingness-to-pay of the downstream firms, her decision depends upon how much downstream profit the licensing scheme can create. With the model developed above, I showed that non-exclusive licensing is preferred under price competition in almost all cases, particularly when the degree of innovation is substantial. Licensing through a royalty scheme tends to increase competition between firms by creating asymmetrical returns for high- and low-quality firms, whereas licensing through a fixed fee tends to moderate competition and creates a market with only high-quality products. I find that two-part tariffs are never optimal because it is impossible for the innovator to facilitate a downstream equilibrium in which either firm benefits. In general, which of the two effects--generating market volume or relaxing downstream competition---dominates depends on the specific parameterization of demand and the extent of the innovation.

For each strategy, the innovator's profit depends on the degree of innovation (λ). Allowing net license revenue to vary with the degree of innovation sheds some light on how potential licensing revenue changes if the extent of innovation they are tasked with marketing varies. By calculating innovator profit over a range of λ values, I find two critical values for λ. There are two effects involved: First, lower values of λ imply more vertical differentiation and higher profits for the high-quality good. This is the "quality-improvement" effect. Second, lower values of λ also imply greater asymmetry in returns and a lower volume-enhancing effect. When the degree of innovation is relatively small ($0.25 < \lambda < 1$), the quality-improvement effect is dominated by the volume-increasing effect, and the innovator does not make as much profit. When the innovation is relatively large ($0.15 < \lambda < 0.25$), higher quality begins to dominate the volume effect, and the outcome for the innovator improves. When the innovation is very large ($0 < \lambda < 0.15$) the innovator is almost certain to make a large profit. Therefore, when developing and licensing a new technology, university TTO administrators should be aware of the existence of this "threshold effect" when determining an optimal licensing strategy. More specifically, I show that for our specific parameterization, there are two such thresholds: The first, at $\lambda = 0.25$, guarantees a positive profit, and the second, at $\lambda = 0.15$, offers the promise an even larger profit. The exact values of these thresholds will clearly depend on the nature of the product and the existing competitive structure, but I provide at least theoretical evidence that they are likely to exist.

In order to demonstrate which of the two effects shown above dominates over a reasonable parameterization of the model, I provide a numerical simulation of the net license revenue attainable by the innovator under a range of possible λ values. In table 2.1 below, I illustrate the relationship between innovator profit and the magnitude of the innovation under each strategy. To keep the simulation as "clean" an experiment as possible, I fix b at a moderate level of $b = 0.50$ and consider the following levels of λ: $\lambda = 0.50$, $\lambda = 0.25$, $\lambda = 0.15$, and $\lambda = 0.05$. The results in table 1 show that, under each licensing strategy, the patent holder's profit increases as the extent of innovation becomes larger. When innovation is sufficiently large ($\lambda = 0.15$, $\lambda = 0.05$), profit is substantially higher than when the innovation is relatively small ($\lambda = 0.25$, $\lambda = 0.50$). Overall, however, this experiment shows that the preferred strategy is a non-exclusive fixed fee. The potential profit under this preferred strategy is followed closely by a non-exclusive royalty. At least for the range of parameters that are reasonable for our problem, therefore, it appears as though the competitive exclusion of low-quality products is a desirable outcome from the perspective of the innovator.

Table 2.1: Comparison of Innovator Profits and the Extent of Innovation

		Fixed Fee	Royalty	Two-Part
b=0.50, λ=0.50	Exclusive	−0.14	−1.93	N.A.
	Non-Exclusive	1.97	−0.34	N.A.
b=0.50, λ=0.25	Exclusive	49	17	N.A.
	Non-Exclusive	97	45	N.A.
b=0.50, λ=0.15	Exclusive	458	209	N.A.
	Non-Exclusive	873	433	N.A.
b=0.50, λ=0.05	Exclusive	39,850	19,788	N.A.
	Non-Exclusive	77,636	39,118	N.A.

Note: N.A. indicates that licensing will not occur as doing so is in neither downstream firm's interest.

The extent of innovation is clearly important to the potential for innovator profits. However, I also maintain throughout that horizontal differentiation is also likely to influence the amount of revenue innovators can earn from licenses. I examine the horizontal differentiation effect by allowing the *b* parameter to vary and calculate a range of innovator profits over a range of λ values. These results are shown in table 2.2. When selling an exclusive license, I find that the more innovator profits rise the more substitutable are the products downstream. This is because firm 2 is able to draw consumers more easily from the low-quality market and the innovator benefits accordingly, both when royalties and fixed fees are used. On the other hand, innovator profit falls in the degree of substitutability when licenses are sold on a non-exclusive basis. When products from the two firms are not substitutable, I have the usual horizontal differentiation effect: Each enjoys a measure of local monopoly power and earns higher margins as a result. Because both purchase a license, the innovator is able to extract more profit from them, whether through a fixed fee or through a royalty scheme.

Table 2.2: Comparison of Profits and Horizontal Differentiation

		Fixed Fee	Royalty	Two-Part
b=0.90, λ=0.10	Exclusive	2,506	1,230	N.A.
	Non-Exclusive	4,477	2,301	N.A.
b=0.60, λ=0.10	Exclusive	2,444	1,104	N.A.
	Non-Exclusive	4,612	2,331	N.A.
b=0.30, λ=0.10	Exclusive	2,410	1,158	N.A.
	Non-Exclusive	4,752	2,364	N.A.
b=0.10, λ=0.10	Exclusive	2,400	1,150	N.A.
	Non-Exclusive	4,850	2,387	N.A.

Note: N.A. indicates that licensing will not occur as doing so is in neither downstream firm's interest.

2.5 Discussion.

In this chapter I study the optimal licensing strategies under price competition in a duopoly scenario. The results suggest different marketing implications with various strategies. That is, licensing through a fixed fee (the innovator) is able to extract the licensing firms' increased profits, but is not able to control industry output; where as licensing through a royalty, the innovator is able to manipulate the cost structure of the licensing firms, which provides a measure of control over the final output. Two-part tariffs have the potential to generate the most revenue, but licenses will never be purchased this way due to the lack of profitability to the licensed firm(s). Moreover, there are "innovation thresholds" beyond which potential license revenue is likely to be significantly greater than if the innovation were less drastic. Because there is a greater reward to the innovator's institution if the innovation is large, and firms can still differentiate downstream, research officers should encourage "bold" innovations if they want to maximize returns from their portfolio of research.

This research has some limitations. In the model I consider two downstream firms with an outside innovator, which may not always be the case in the horticultural industry. Future research should extend our framework to study the optimal licensing strategies when the innovator is an incumbent and when there are more than two players in the downstream market.

CHAPTER 3. ESSAY 2: VARIETY ADOPTION AMONG FARMERS IN A SOCIAL NETWORK

3.1 Introduction

Variety adoption has long been of interest to agricultural economists. Griliches' (1957, 1960) research on hybrid corn in the USA was not only one of the first economic studies of adoption of rural innovations by economists, but helped establish agricultural economics as an important, independent field of study (Griliches, 1957, 1960). Since then, many theoretical and empirical studies have focused on adoption and diffusion of new technologies and new plant varieties (Feder, Foster and Rosenzweig 1995; Conley and Udry 2001; Sunding and Zilberman 2001). Because technology adoption is a necessary precondition for broader productivity improvement and economic development, this line of research seeks to examine, and perhaps to suggest remedies for, the barriers to adoption. I am interested in one specific barrier of new variety adoption, namely the lack of information, in the context of maize farming in Mozambique.

Mozambique is a country with excellent agricultural conditions, enabling the cultivation of a great variety of crops such as maize, sorghum, and many types of fruits (Mission Report 2007). Yes, despite the many advantages to doing so, the rate of adopting new, improved maize varieties remains relatively low (11% in 2011). Although farmers who have adopted improved maize varieties are distributed throughout all regions and provinces of the country, adopters and non-adopters appear to cluster geographically (Figure 3.1). This pattern raises questions as to the nature of the relationships among farmers that may either aid adoption, in the case of adopting clusters, or hinder adoption where non-adoption appears to be the norm. Theories for non-adoption range from technological barriers to policy limitations, but patterns of clustering among neighbors seem to suggest a strong informational component. In fact, information is found to be the most important factor in rural development (Bajeree 1973, Bikhchandani et al. 1992, Case 1993) so this pattern would be consistent with previous research on adoption. In this chapter, I explain how the spatial patterns of households influence adoption through information transition.

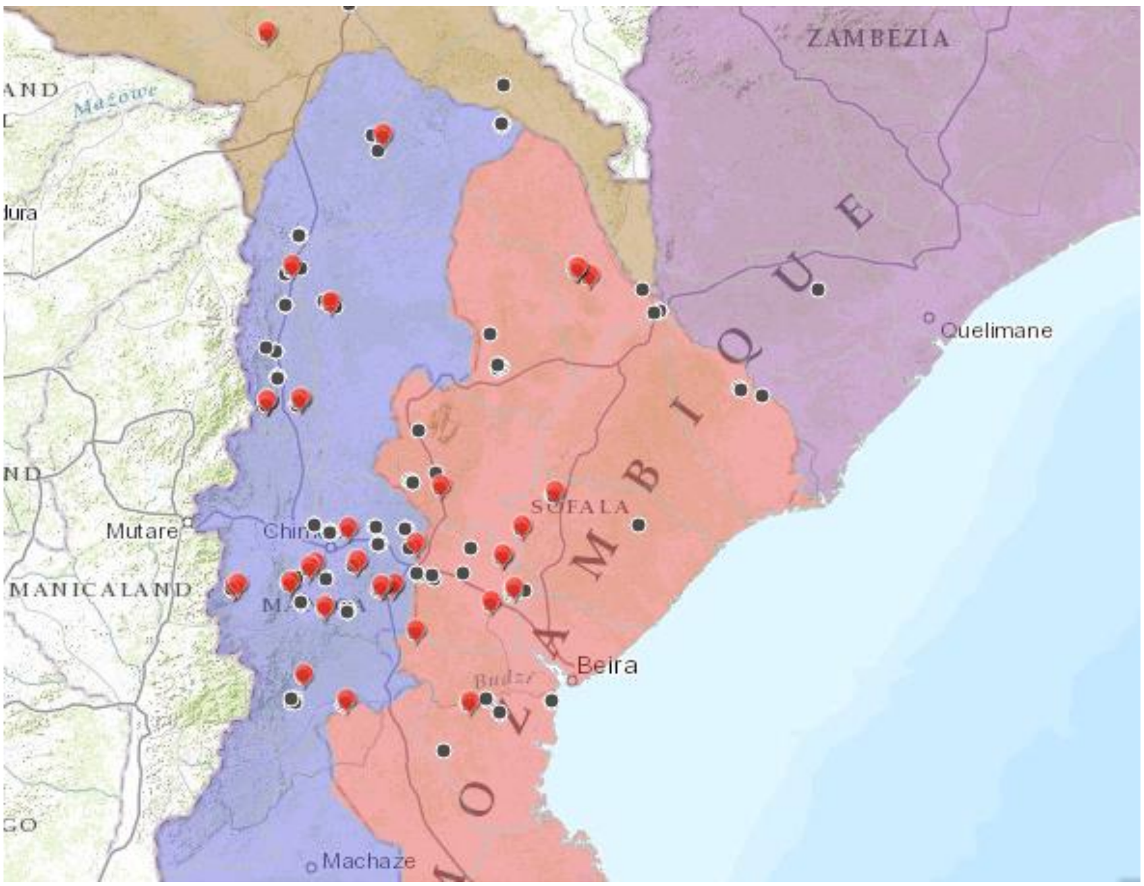

Figure 3.1. Sampled Household in Sofala and Manica with Adopters (red) and Non-adopters (black).

Information can be acquired in any one of a number of ways: through formal education, social media, trade organizations, or simply by agents imitating others. Perhaps most important in developing countries, where information sources are largely informal, is the acquisition of information from neighbors and other farmers. The process of acquiring information in this way is often referred to as "social learning" where a small number of "leaders" adopt, and the rest observe before taking action[10]. Social learning is endogenous as it describes the process through which farmers learn from each other, where non-social learning is, to some degree, constrained by availability of resources such as household income, policy and government subsidies. In this chapter I study both types of learning and their influences on new variety adoption as a means of explaining the clustering phenomenon observed in Figure 3.1.

Relying mostly on micro-level data, studies consistently find both social learning and non-social learning to be significant (Bikhchandani et al. 1992, Foster & Rosenzweig, 1995, Abdulai and Huffman 2005, Bandiera and Rasul 2006). Foster and Rosenzweig (1995) study the adoption of high yielding seed varieties (HYVs) in India and identify the barrier to adoption to be imperfect knowledge about the management of the new seeds. They find that learning by a farmer's own experience, and by observing neighbors, could increase adoption. Moreover, individuals don't consider the welfare of the community as part of the return. Similarly, Abdulai and Huffman (2005) examine farmer's adoption of crossbred technology of cows in Tanzania and finds that adoption depended positively on the proximity of his farm to other users, education, and on his access to credit and contact with extension agents. With a focus on social learning, Bandiera and Rasul (2006) consider a farmer's network of family and friends in the adoption of sunflower in Northern Mozambique and find that there is a quadratic

[10] To make a clear distinction between social learning and the other sources of information, I define information acquired from sources other than social learning (such as education, social media, extension service and organizations) as non-social learning.

relationship between positive information externalities and the number of adopters in one's network. However, they use a purely temporal adoption model where the mechanism is one-dimensional. That is, they assume that social influence operates as a function of the number of adopters in the farmer's network. This assumption is unrealistic because social learning works through interactions within network, so is inherently spatial in nature.

Spatial econometric analysis is a viable tool to assess the adoption decision in a multi-dimensional way. Spatial relationships between agents are captured through "spatial weight matrices" which measure the distances between each agent and each other agent in the network. In my example, spatial weight matrices capture the relational proximity between two farmers and reflect such inter-relational characteristics. I measure the spatial relationship between farmers with three types of matrices: immediate neighbor, geographical distance, and, extended neighbors (*rook contiguity*). As expected, immediate neighbors play an important role in adoption decisions. Moreover, I find that distance does not measure the relationship between farmers as those who live in close geographical proximity do not influence each other significantly. Instead, farmers exchange information with extended neighbors, regardless of the distance.

Prior studies in this area also ignore the cost of learning. The cost of learning is critical because, in reality, information cannot be acquired for free. Assuming information is costless is not benign as the outcomes of a costly-learning model are fundamentally different in that the relevant pure strategy to the adoption game with imperfect information differs solely due to the structure of costs. For example, by imposing a cost of waiting and assuming a difference in the quality of information possessed by different farmers, Zhang (1997) shows theoretically that there is an initial delay before the first adopter uses the technology. Moreover, farmers are subject to unique benefits and costs due to the quality of information they receive.

I contribute to the literature on the adoption of new agricultural technologies in a number of ways. First, I include the cost of learning in a theoretical model of adoption, and show that, by doing so, the conclusions regarding the nature of strategic delay are fundamentally altered. Namely, I show that a delay is optimal, or strategic, if farmers do not have accurate information regarding the new variety, and cost works against delay because the more a farmer waits the more costly it will be. Initial information comes from non-social learning such as education, extension services, social media and organizations. Non-social learning is positively related to the accuracy of information, therefore, when abundant, non-social learning is available to a farmer, he will adopt early in order to grasp the "first-mover advantage." However, when non-social learning is unavailable, farmers will depend on social learning instead, which results in a strategic delay. In addition, I show that strategic delay is a trade off between the costs and benefits of learning. On one hand, being connected to other farmers will generate information externalities that benefit the decision making process. On the other hand, forming links in a network is costly (Acemoglu, Bimpikis, Ozdaglar 2009). Therefore, the adoption decision is complicated by a farmer's network, through which he engages in social learning.

Second, I test the implications of my model using empirical data that describes the adoption of new maize varieties by farmers in Mozambique. My model is inherently spatial, drawing on the isomorphic relationship between geographic and social space noted by others (Anselin, 2002, Kalnis 2003, Lesage 2009), so information flows throughout the network in a multi-dimensional way. I detect social learning in the model. Social learning is tested through a “spatial lag parameter,” which measures the extent of spatial relationships among proximal agents. I find significant social learning among farmers. More specifically, a farmer observes from his nearest neighbors and bases his adoption decision on theirs. Moreover, geographic proximity does not contribute significantly to learning as immediate and extended households do, independent of how close they are. When geographical distance is critical, then the policy implication is clear -- improve the infrastructure such as roads and public transportation. In this case where geographical distance is not crucial, government should explore alternative venues of information distribution such as increasing cell phone and Internet use.

Besides learning from others, learning from experience is negatively related to adoption, as I find that farmers who have planted the new variety before are not likely to plant again. This is counter-intuitive, as farmers who have adopted should stay with the variety because the new variety is an improvement from the traditional variety. Combined with the fact that most adopters produce for sale instead of own consumption, this notion may suggest a significant difference in tastes between the old and new varieties. From the policy standpoint, the government should focus on retaining existing adopters as well as recruiting new adopters. A number of suggestions include improving the traits of the new variety such as taste, distributing agriculture credits to lower the cost (risk) of switching, and build an efficient demand-supply market for maize to encourage economies of scale.

Non-social learning plays an important role in strategic delay too. Initial knowledge acquired from education and extension helps the farmer to evaluate the new variety. A more informed farmer is more likely to adopt than a less informed farmer. Training tailored specific to the new variety is extremely helpful in promoting adoption because it provides direct information on the new variety. Extension is also found to be significant and positively related to adoption despite the limited accessibility to farmers. For example, only 2.2% of the farmers received training on planting the new variety. Even so, the training turns out to be significant and positively related to adoption. Given that network is an important tool of information exchange, the policy implication here is to educate influencers in the local networks and let them pass along the knowledge.

The chapter is organized as follows. In the next section I propose a theoretical model that describes the adoption game between farmers, the primary output of which is a rationalization of strategic delay that is unique to the literature. In the third section, I test the implications of my theoretical model using farm-level data from maize-variety adoption in Mozambique. A fourth section summarizes the empirical results, and provides a discussion of the policy implications of my findings. The final section concludes, and offers some broader implications for the effect of social learning on strategic delay in contexts other than my specific empirical example.

3.2 Theoretical Model

Social learning assumes that rational, profit-maximizing agents respond to information generated by other adopters (Banerjee 1992, Caplin and Leahy 1998, and Chamley and Gale 1994). With social learning, farmers decide whether or not to adopt, or when to adopt a new variety, by balancing the marginal benefits and costs of using a new technology or variety. I this section, I first introduce a general model to illustrate how adoption is achieved through a pure Bayesian equilibrium, followed by a more specific theoretical model that frames the empirical example in the next section.

Assume there is a set of farmers *I*, with a set of actions (pure strategies) for each player *i*: s_i. Farmers are hetergenous in that each farmer is of a type $\theta_i \in \Theta_i$. The payoff function for each player *i* is $u_i(s_1, \ldots, s_I, \theta_1, \ldots, \theta_I)$. There exists a (joint) probability distribution p $(\theta_1, \ldots, \theta_I)$ over types$(\theta_1, \ldots, \theta_I)$ of inputs. Farmers do not observe the types of others but instead make inferences based on signals. Importantly, the strategy spaces, payoff functions, possible types, and the prior probability distribution are assumed to be common knowledge. A (pure) strategy for player *i* is a map s_i: $\Theta_i \rightarrow \mathrm{S}_i$, that prescribes an action for each possible type of player *i*. Assuming layer types are drawn from some prior probability distribution p$(\theta_1, \ldots, \theta_I)$, Given p$(\theta_1, \ldots, \theta_I)$ I compute the conditional distribution p$(\theta_{-i}| \theta_i)$ using Bayes rule[11]. Player i knows her own type and evaluates her expected payoffs according to the conditional (posterior) distribution p$(\theta_{-i}| \theta_i)$, where θ_{-i} = $(\theta_1,, \ldots \theta_{i-1},\theta_{i+1},, \ldots, \theta_I)$. A farmers' welfare is measured by the expected payoff from adopting the new variety. Assuming the payoff functions, possible types, and the prior probability distribution are known, the expected payoffs of player *i* of type θ_i is:

$$U(s_i', s_{-1}(\cdot), \theta_i) = \sum_{\theta_i} \mathrm{p}(\theta_{-i} \mid \theta_i) u_i(s_i', s_{-i}(\theta_{-i}), \theta_i, \theta_{-i}). \qquad (3.1),$$

which is a function of expected payoff after observing others.

[11] An illustration of Bayesian updating is provided in appendix B.

The strategy profile s(·) is a (pure strategy) Bayesian Nash equilibrium if for all $i \in I$ and for all $\theta_i \in \Theta_i$, and is written as:

$$s_i(\theta_i) \in \arg max \sum_{\theta_i} \mathrm{p}(\theta_{-i} \mid \theta_i) u_i(s_i', s_{-i}(\theta_{-i}), \theta_i, \theta_{-i}). \qquad (3.2),$$

where the strategy profile is a set of strategies that maximizes individual's expected payoff.

Consider a Bayesian game with continuous strategy spaces and continuous types. If strategy sets and type sets are compact, payoff functions are continuous and concave in own strategies, then a pure strategy Bayesian Nash equilibrium exists (Acemoglu, Bimpikis and Ozdaglar 2014). Hence, a Bayesian Nash equilibrium is a Nash equilibrium of the expanded game in which each player i's space of pure strategies is the set of maps from $\Theta_i \to \mathrm{S}_i$.

Now consider a sequence of farmers (i = 1, 2, ...I) making decisions $x_i \in \{0, 1\}$ where $x_i = 1$ means farmer i decided to adopt; and $x_i = 0$ means farmer i decides to wait. Farmer i has a neighborhood $B(i) \subseteq \{1, 2, ..., \mathrm{I} - 1\}$ and observes the decisions x_I for all $i \in$ B(I). The set *B(I)* is private information but each farmer has an iid private signal σ_i. The signal is generated according to distribution $\mathbb{F}$. Farmer i's information set is $K_i = \{\sigma_i$, B(I), x_i for all k $\in$ B(I)$\}$. A strategy profile is a sequence of strategies $s_i \in \{S\}_{i \in I}$, with a probability distribution P_s over $\{x_i\}_{i \in I}$. A strategy profile S is a pure-strategy Perfect Bayesian Equilibrium if for all i:

$$S_i^*(K_i) \in \arg max \, \mathrm{P}_s(x = s|K) \qquad (3.3),$$

which means that best strategy depends on the probability of taking an action given available information.

The adoption decision of farmer *i* is then:

$$x_i \begin{cases} 1, if\ \mathrm{P}_s(s = 1|\sigma_i) + \mathrm{P}_s(s = 1|\mathrm{B(I)}) > 1 \\ 0, if\ \mathrm{P}_s(s = 1|\sigma_i) + \mathrm{P}_s(s = 1|\mathrm{B(I)}) < 1 \end{cases} \quad (3.4),$$

where $\mathrm{P}_s(s = 1|\sigma_i)$ is the private belief about the new variety and $\mathrm{P}_s(s = 1|\mathrm{B(I)})$ is the social belief about the variety. This means that when the probability of getting positive payoff (if adopting the variety) is confirmed by both private information (non-social learning) and social learning, the farmer will choose to adopt. Otherwise the farmer will choose to wait.

Following Acemoglu, Bimpikis, Ozdaglar (2009), forming links in a network is costly, so let c_{ij}^I denote the cost incurred to maintain the link between farmer *i* and *j* over a communication network of G^i. Each farmer receives a private signal σ_i~ $\mathbb{F}$ from direct neighbors, and makes a decision of whether to adopt (1) or wait (0) at time *t*. Assuming δ (δ < 1) is the discount factor for waiting, π is the payoff from adoption, and τ is the time taken up to the action; the payoff function is:

$$u_i(x_i^I, s_i) = \begin{cases} \delta^\tau \pi - \sum_i c_{ij}^I, if x_{i,\tau}^I = 1\ and\ x_{i,t}^I = 0\ for\ t < \tau \\ 0, otherwise \end{cases} \quad (3.5),$$

where the utility from adopting or not adopting depends on the profit of adopting (discounted by time) and cost of learning. When the profit trumps cost of learning, a farmer will adopt. When the profit is not sufficient to cover the cost of learning, a farmer will choose to wait.

A strategy profile ${s_i}^{I*}$ is a Perfect-Bayesian Equilibrium if for all i and t

$$s_{i,t}{}^{I*} \in \arg \max_{y \in \{0,1\}} E_{(y, s_{-i,t}{}^{I*})}\, u_i((x_i^I, s_i)|K_i) \quad (3.6)$$

Where the optimal strategy is now derived from the maximized expected payoff under private and public information.

Intuitively, the equilibrium point of adoption can be found by maximizing the expected utility function that considers the payoff and cost of information exchange in a network, given that farmers are subject to both private observational information and public network information. In this model, information comes from social networks, typically neighboring farmers, as well as non-social sources such as extension services, or traditional media.

In order to add specificity to the model above, I impose a number of assumptions that reflect the reality of adoption in developing countries. Within this general Bayesian framework, I consider a more specific mechanism akin to the target input model of Rosenzweig (1995) and Bandiera and Rasul (2006), while introducing costly network relationships. A target input model assumes that an optimal input is stochastic and unknown. This model is appropriate because optimal input use appears empirically to be central to farmers' concerns regarding adoption, and that its implications are easily tested using appropriate econometric methods.

Others use a similar framework to study adoption (see Foster and Rosenzweig 1995, Conley and Udry1994, Bandiera and Rasul 2006), but none consider the cost of obtaining information. Including costly information is critical, because the optimal strategy alters based on the cost structure.

At the core of the Rosenzweig (1995) model is a Bayesian learning process. A Bayesian updating process implies that the farmer updates his previous knowledge (prior belief) of his optimal input employment based on his own trials and signals about inputs from others' trials. He reaches a new optimal input target, a posterior belief, based on these signals. Bayesian models are relatively common in an adoption context (Feder and O'Mara 1982; Leathers and Smale 1991; Rosenzweig 1995; Ghadim and Pannell 1999; Bandiera and Rasul 2006), and for good reason. Namely, farmers adopt because the expected profit is positive. Absent from existing models, however, is the notion that adoption is costly. Allowing for a cost of adoption, I show that the model can produce substantially different results compared to when cost is ignored. Therefore, I add a cost structure to the Bayesian framework in order to derive a set of implications that are more descriptive of the context of variety adoption in developing countries.

The basic features of this model are as follows. First, individuals decide on the input amount, which is a random variable with a known mean and variance. Second, payoffs are decreasing in the square of the distance between actual input use and the target, so the closer is the farmer's actual input to his target input, the higher his payoff. Third, each individual can observe others' inputs in a previous period, and thereby update their input in the following period in a Bayesian fashion (Rosenzweig 1995).

I follow this framework in that I consider a new maize variety as a target input in which farmers update their knowledge regarding its performance attributes in a Bayesian fashion. I assume each farmer has private information about his[12] idiosyncratic target input value (fertilizer usage, irrigation, sowing techniques, pest control, etc.)—through non-social learning such as training and extension services, but shares public information that is generated by early adopters. However, public information is not the same to all

[12] I use the male pronoun to represent farmers because a majority of farmers are male.

farmers because each farmer's network is different. I then incorporate unique networks into an expected profit function.

For farmer i in period t, output q_{it} declines in the square of the distance between the actual input used k_{it}, and the optimal input target in time t, κ_{it}, while η_0 is the maximum output given the underlying technology. The production function is written as:

$$q_{it} = \eta_0 - (k_{it} - \kappa_{it})^2 \qquad (3.7),$$

where farmers have prior beliefs about the optimal input κ_{it}, and update their belief about the optimal input according to the actual input k_{it}. The target input level is not known at the time the input is chosen and has a mean of κ^*:

$$\kappa_{it} = \kappa^* + \mu_{it} \qquad (3.8).$$

In period t farmer *i* has beliefs about κ^* which are distributed as $N(\kappa_{it}{}^*, \sigma^2_{\kappa_{it}})$. Define μ_{it} as a idiosyncratic shock that is i.i.d with a mean of 0 and a known variance σ^2_μ. As $E(\mu_{it}) = 0$, to maximize expected output farmer i uses the expected optimal target level as his input, so $k_{it} = E_t(\kappa_{it}) = \kappa_{it}{}^*$. Therefore the expected output is:

$$E_t(q_{it}) = \eta_0 - E_t\big(k_{it} - E_t(\kappa_{it})\big)^2 = \eta_0 - \sigma^2_{kit} - \sigma^2_\mu \qquad (3.8)$$ [13]

Which is a function of the optimal input, variation from initial knowledge and variation from learning.

The expected output is a function of maximum output possible given a farmer's capacity less the deviation caused by his inaccurate input estimate. Repeating the process *N* (*N*=*t-1*) times, the farmer updates his knowledge regarding the optimal target input

[13] These derivations follow Rosenzweig (1995).

from his initial belief in period 1 to period t based on the signals generated by other people's inputs. Applying Bayes's rule, the posterior belief [14] is:

$$\sigma^2_{k,it} = \frac{1}{\frac{1}{\sigma^2_{ki,0}}+N\frac{1}{\sigma^2_{\mu}}} \qquad (3.9),$$

where $\rho_0 = \frac{1}{\sigma^2_{ki0}}$ is the precision of the information generated by i's own experiences prior to the updating process, and $\rho_L = \frac{1}{\sigma^2_{\mu}}$ is the precision of the information obtained each time the process is used, and *N* is the cumulative number of times prior to time t that the process has been employed.

$$\sigma^2_{k,it} = \frac{1}{\rho_0 + N\rho_L} \qquad (3.10).$$

This means that the posterior belief about the new variety is a function of the initial knowledge of the variety, number of trials a farmer experienced, and the precision of knowledge a farmer learns from his network. Substituting this expression back into the expected output yields an expression that includes the optimal input, precision of initial knowledge, number of trial, precision of knowledge learned, and is written as:

$$E_t(q_{it}) = \eta_0 - \frac{1}{\rho_0 + N\rho_L} - \frac{1}{\rho_L} \qquad (3.11)$$

Costs come from two sources: input costs and the cost of obtaining information. I assume there is costant unit cost, θ_0, to each unit of q_{it} invested. That is, for an investment of q_{it}, the cost is $\theta_0 q_{it}$. More importantly, there are two ways in which precision may influence the cost of obtaining information: First, the precision of knowledge about the new variety prior to adoption, and, second, the precision of

[14] An illustration of Bayesian updating can be found in Appendix B.

knowledge after N repeated observations. Therefore, cost is a function of the number of times observed N, the precision of initial knowledge and the precision of knowledge after observation: $\mathrm{C} = \mathrm{C}\ (\mathrm{N}, \rho_{\mathrm{L}}, \rho_0)$, where $C_N' > 0$, $C_{\rho_0}' < 0$, and $C_{\rho_{\mathrm{L}}}' > 0$. That is, cost increases with the number of times observed, decreases with the precision of knowledge and increase with learning.

More precisely, I assume there is cost involved with waiting so that: $C_N' > 0$. As the number of trials a farmer experiences before he adopts, the opportunity cost increases. The notion that $C_{\rho_0}' < 0$ suggests that the more knowledge a farmer has about the new variety prior to adoption, the less costly it will be to make a prior judgment about the optimal input value. Such prior knowledge can be obtained through non-social learning such as education as to the importance of new varieties, receiving extension services, and collecting information from non-social learning such as radio, TV and literacy meetings organized by trade associations. I represent this idea by allowing the initial precision of knowledge to be a function of non-social learning or: $\rho_0 = \rho_0(Education, Extension, Media)$, where Education indicates the years of schooling a farmer has received, Extension indicates the efforts to attend extension services, and Media indicates the receipt of information from social media. Obtaining information from any of these sources causes the farmer to incur a positive cost. Optimality requires that the farmer equate the marginal cost of new information with the marginal benefits. Because farmers can make more precise prior judgments about the new variety ($\rho_0' > 0$) the marginal benefits increase accordingly.

The cost of acquiring information plays a critical role in my model. Social networks are formed as the result of individual decisions that trade off the costs of forming and maintaining links against the potential rewards from doing so (Scott 2010). Having connections with others helps a farmer update his knowledge about the new variety. However, direct links are costly (Bala and Goyal 2000) because of the effort required to form and maintain the connections. Particularly in the context of rural areas of developing countries, infrastructure such as roads and markets are under-developed, which imposes additional constraints on farmers' abilities to observe their neighbors. Therefore, the spillover effects from observing previous adopters might be limited. More specifically, a farmer will need to consider both the benefits of belonging to a network as well as the cost of maintain such a network when making the adoption decision.

Denote the number of a farmer's direct connections at time t to be $n_i^d(i)_{t-1}$, and the precision of learning is a functions of direct connections: $\rho_L(n^d(i)_{t-1})$, and assume that ρ_L is concave and twice differentiable with $\rho_L'(\quad) > 0$ in $n^d(i)_{t-1}$ and $\rho_L''(\quad) < 0$ in $n^d(i)_{t-1}$. This is consistent with the findings of Bandiera and Rasul (2006) in that as the number of adopters increases, adoption increases initially but falls at the margin due to "information overload".

Profit is defined in terms of expected returns, input costs and the cost of information. Assuming a constant return to output *p*, the profit function is then:

$$\begin{aligned}&\pi_t[p, q_{it}, C]\\&= (p - \theta_0)(\eta_0 - \frac{1}{\rho_0 + N\rho_L} - \frac{1}{\rho_L})\\&- C(N, \rho_L(n^d(i)_{t-1}), \rho_0(Education, Extension, Media))\end{aligned} \qquad (3.12),$$

Where profit is a function of non-social learning, social learning and cost of learning. I derive specific hypothesis in the next section based on this profit function.

3.3 Hypotheses

To shed light on the effect of learning on variety adoption, I derive a number of hypotheses regarding the propensity to adopt based on the profit (3.12) that is a function of network, social and non-social learning, and cost of learning. While intuition suggests that adoption choices should be positively related within networks, theories of social learning indicate that the sign of the relationship is actually ambiguous. On the one hand, the benefit of adopting in the current period is higher when there are many adopters in the network because of the information they provide. On the other hand, having many adopters in the network increases incentives to delay adoption strategically and free ride on the knowledge accumulated by others. If strategic delay considerations prevail, a farmers' propensity to adopt decreases as the number of adopters among his network increase.

HYPOTHESIS 1. *Experienced farmers with precise knowledge of the new variety tend to be "first movers" with greater output, while less experienced farmers will strategically delay adoption until the marginal benefit of learning externalities is equal to the marginal cost of waiting.*

PROOF: When farmers choose to wait for others to adopt, postponing their own adoption, the time lag is known as a strategic delay. The delay is strategic as it is not driven by a failure on the farmer's part, or even a rational response to the real option value embedded in a new technology (Richards and Green 2003), but rather an optimal response to information acquisition in an uncertain world. Several factors can justify a farmer's delaying adoption. When farmers can observe each other's signals for free, delays in decision-making can indeed be rational (Banerjee 1992). Moreover, the length of the delay is sensitive to the reaction speed of each player and the number of players

(Chamley and Gale 1994) because when farmers react quickly to other people's adoption the delay will be shortened. By the same token, when there are many farmers in a network, they produce more information externalities, which will help accelerate the decision process.

In general, first-movers have an advantage but delay occurs due to the risk of making what turns out to be an ex-post poor adoption decision (Farrell and Saloner 1985, Farrell 1987, Farrell and Saloner 1988, Bolton and Farrell 1990, Farrell 1993). Information from others is critical as it changes the precision of the information available. To determine how the precision of information influences expected output, I derive the marginal effect of precision on the expectation of output, where precision is acquired from the farmer's own experience with the variety and the precision of the farmer's initial knowledge of the variety. The marginal effect of information on output is.

$$\frac{\partial \mathrm{E_t}(\mathrm{q_{it'}})}{\partial \rho_0} = \frac{1}{(\rho_{i0} + \mathrm{N}\rho_{\mathrm{L}})^2} > 0 \tag{3.13}$$

While the marginal effect of learning from others is:

$$\frac{\partial \mathrm{E_t}(\mathrm{q_{it}})}{\partial \rho_{\mathrm{L}}} = \frac{\mathrm{N}}{(\rho_0 + \mathrm{N}\rho_{\mathrm{L}})^2} > 0 \tag{3.14}$$

Therefore, both the marginal effect of precision on the farmer's own experience and learning from others are positive. This shows that if a farmer has precise initial judgments about the optimal input, then the expected payoff will increase accordingly. The advantage of such initial judgments diminishes as the number of periods increase, that is, more experienced farmers tends to be the "first movers" because the longer the farmer waits, the lower the profit from adoption. For the farmer who does not possess such precise initial judgments, his expected output will increase as he observes from others though the learning process. In other words, positive externalities derive from repeatedly observing others.

Moreover, farmers learn by their own experiences and observing their neighbors, and actively respond to future information from other adopters (Besley and Case 1997). Munchi (2003) considers the adoption of rice varieties in India and finds that heterogeneity across rice-growing regions, besides farm characteristics, contributes to adoption. Heterogeneity in his context includes accessibility to extension service, formation of local network and education. Abdulai and Huffman (2005), meanwhile, show that farmers' adoption of crossbred cattle in Tanzania depends positively on the proximity of his farm to other users, education, and on his access to credit and contact with extension agents. Therefore, education, extension services and social media are also likely to be important factors in explaining expected payoff.

HYPOTHESIS 2. *Non-social learning is positively related to adoption.*

PROOF: The proof of Hypothesis 2 depends on the comparative statics of the equilibrium above with respect to education. Non-social learning such as education increases expected profit two ways. First, it increases expected output by helping farmers better estimate optimal input levels. By increasing returns through precision, output rises. Second, education reduces the learning related to make an initial decision, which causes profit to rise as:

$$\frac{\partial \pi_t}{\partial Education} = (p - \theta_0)\left(\frac{1}{\rho_0 + N\rho_L}\right)^2 \frac{\partial \rho_0}{\partial Education} - \frac{\partial C}{\rho_0}\frac{\partial \rho_0}{\partial Education} > 0 \tag{3.15}$$

The expression is positive because $\frac{\partial C}{\rho_0} < 0$ and $\frac{\partial \rho_0}{\partial Education} > 0$, and the first part is positive if incremental returns exceed the marginal input cost. This shows that non-social learning increases the expected payoff for the farmer.

Other than education, extension service and social media, a farmer's network represents his source of private information. In particular, it is the size of the network that conveys the most information about the extent of social learning.

HYPOTHESIS 3. *The number of adopters in a network has an ambiguous effect on output.*

PROOF: The comparative static effect of a famer's direct connections on his expected payoff is found by again finding the marginal effect of network size on equilibrium profit. Profit as a function of network size is given by:

$$\pi_t[p, q_{it}, C] = (p - \theta_0)(\eta_0 - \frac{1}{\rho_0 + \mathrm{N}\rho_\mathrm{L}(\, n^d(i)_{t-1})} - \frac{1}{\rho_\mathrm{L}(\, n^d(i)_{t-1})}) - C(\mathrm{N}, \rho_\mathrm{L}(\, n^d(i)_{t-1}), \rho_0\,) \qquad (3.16)$$

Where $n^d(i)_{t-1}$ is the number of direct links in a farmer's network. The comparative static with respect to network size is:

$$\frac{\partial \pi_t[p, q_{it}, C]}{\partial\, n^d(i)_{t-1}} = (p - \theta_0)(\frac{1}{(\rho_0 + \mathrm{N}\rho_\mathrm{L})}^2 + \frac{1}{n^d(i)_{t-1}{}^2})\frac{\partial \rho_\mathrm{L}}{\partial\, n^d(i)_{t-1}} - \frac{\partial C}{\partial \rho_\mathrm{L}}\frac{\partial \rho_\mathrm{L}}{\partial\, n^d(i)_{t-1}} \qquad (3.17),$$

where the first part represents the positive externality associated with having more adopters in the network. That is, the more adopters there are in the network the more the farmer is able benefit from the learning process. The second part is interpreted as the cost associated with maintaining such a network. Network costs can be associated with travel expenses to the nearest adopter, tradeoff in information, necessary social norms or gift giving. There are mixed effects on the incentive to adopt as the number of adopters in a network increases.

$$\frac{\partial \pi_t[p, q_{it}, C]}{\partial\, n^d(i)_{t-1} \partial N} = -2(p - \theta_0) \frac{\partial \rho_L}{\partial\, n^d(i)_{t-1}} \frac{1}{(\rho_0 + N\rho_L)^3} < 0 \qquad (3.18)$$

This derivative is negative because $\frac{\partial \rho_L}{\partial\, n^d(i)_{t-1}} > 0$, or having direct links to adopters provides additional information on the new variety, which in turn will increase the farmer's expected payoff. In this sense, waiting is not a preferred strategy. On one hand, the number of adopters in a network increases as the farmer waits, but the marginal advantage from learning diminishes the longer the farmer waits. Therefore, the timing of adoption depends on the number of adopters, the private cost of learning, and the farmer's prior knowledge regarding the new variety.

The timing of adoption is a critical concern when market-entry is strategic. In a game-theoretic framework, players maximize their utilities with regard to private information they own and public signals released by others. In aggregate, the micro-foundations of adoption lead to an aggregate adoption curve Zhao (2001) studies an infinite period game with finite number of players and incomplete information and finds that the adoption rate differed in different stages of diffusion. Similarly, Kapur (1995) studies adoption in a complete information dynamic game of identical agents and finds that agents randomize the timing of adoption and can end up adopting at different times. .

HYPOTHESIS 4:*Famers adopt at different times due to individual cost of obtaining information.*

PROOF. In my framework, waiting cost is also important in determining the speed of adoption. If there is no cost of waiting, then the “followers” will wait indefinitely because more information is generated. However because waiting is costly, the follower will adopt at time period t where t $|(\frac{\mathrm{N}}{(\rho_0+\mathrm{N}\rho_\mathrm{L})^2} - C_N') = 0$, that is, when the marginal benefit from waiting is equal to the marginal cost of waiting. Because individual waiting costs differ, due to heterogeneity in signal precision, farmers will adopt at different times.

Clearly, the benefits generated by network externalities decrease as the number of adopters increases. So, the size of an efficient network, where efficiency refers to an optimal trade off between the benefit from network externalities and the cost of marinating such a network, depends on the relative marginal effects of size on cost and returns. That is, if the network expands by including one more connection, it will not generate additional learning externalities.

HYPOTHESIS 5. *The efficient network is not infinite due to the cost of network formation.*

Setting the marginal profit with respect to network membership to zero and solving for *n* finds the efficient network size:

$$\frac{\partial \pi_t[p,q_{it},C]}{\partial\, n^d(i)_{t-1}} = (p-\theta_0)\left(\frac{1}{(\rho_0+\mathrm{N}\rho_\mathrm{L})}^2 + \frac{1}{n^d(i)_{t-1}{}^2}\right)\frac{\partial \rho_\mathrm{L}}{\partial\, n^d(i)_{t-1}} - \frac{\partial C}{\partial \rho_\mathrm{L}}\frac{\partial \rho_\mathrm{L}}{\partial\, n^d(i)_{t-1}} = 0, \tag{3.19}$$

which implies that an efficient network is a function of the initial knowledge of the variety, learning from others, number of repeated trials, the number of direct links a farmer has, and the cost of maintaining such direct links. In reality, the efficient network size differs based on many factors such as geographical proximity, closeness of society and the availability of transportation and infrastructure.

In summary, this section shows how farmers can optimally delay adopting a new variety – a strategic delay. More specifically, the incentive to delay adoption increases in the number of other adopters because use of the new variety by network members creates more useful knowledge. However, the value of information a farmer receives from his own adoption is lower as more network members adopt. Therefore, the gain in future profitability from an additional trial to the farmer with the new technology is decreasing in the number of trials from all network members. Moreover, because it is costly to observe information and maintain a network, the adoption decision and timing of adoption is now complicated by the network structure experienced by the individual farmer. Intuitively, a farmer will maintain an efficient network size where he can observe public information from others while minimizing the cost of using the network.

Empirically testing the hypothesis in a structural way is not possible due to data constraints. However, in the next section I describe an empirical example that allows me to test the implications of my theoretical model in a straightforward way.

3.4 Empirical Study: New Maize Variety Adoption in Mozambique

Agriculture plays a crucial role in the lives of Mozambicans. The agricultural sector provided employment to 81% of the population in 2014, and added $4.08 billion to GDP in 2012 (NationMaster, 2014). Given the dominance of agriculture in the macroeconomy, general economic growth and poverty alleviation in Mozambique are practically impossible to achieve without sustainable development of the agricultural sector. For this reason, improved varieties (e.g., improved maize and beans) have the potential to increase production, as well as increase income and improve the standard of living for farm households.

Unfortunately, the rate of new maize variety adoption in Mozambique remains low. Approximately 11% of agricultural households planted improved maize varieties in 2011, largely because households question the economic profitability of cultivating improved varieties of maize and other staple food crops. In Mozambique, during good years when yields are high, households receive low prices. However, because producer prices are regulated, and do not respond to shortages, the price of grain does not increase in bad years, regardless of low production. On top of that, maize production is often operated off smallholders, who are vulnerable to adverse climatic conditions and natural disasters such as drought and flooding (Mission Report 2007).

In order to encourage households to adopt new agricultural varieties and increase agricultural productivity, the government and non-government organizations have increased the number of extension agents and programs beginning in 2004 (Lopes 2010). However, only 15% of rural households had access to extension services from either the government or nongovernment organizations (Lopes 2010). This suggests that the number of households who are aware of new agricultural technologies and improved crop varieties is limited. As a result, local networks assume a critical role in facilitating the adoption of improved new varieties. Just how important local networks are, however, is an empirical question.

3.4.1 Data

My data consists of farm-level information gathered through an annual survey administered by the Ministry of Agriculture (MINAG). Through this survey, the

Mozambique government samples crop planting at an individual- household level. The resulting dataset includes household-level observations on sociodemographic information, and descriptions of farmers' familial and friend networks, in addition to a comprehensive set of production data. The National Agriculture Survey (Trabalho do Inquérito Agrícola, or TIA), which was first conducted in 1993 by MINAG staff from the Directorate of Economics in collaboration with colleagues from Michigan State University (MSU), employs standards from the National Statistics Institute (INE). The TIAs uses a stratified, clustered sample design that is representative of rural small- and medium-holders at the provincial and national levels.

Smallholders are the backbone of the agricultural sector. A smallholder is defined as having less than 10 hectares of cultivated area, fewer than 10 cattle, 50 goats, pigs or sheep, and 5,000 chickens (TIA Dissemination, 2007). It is estimated that there are over 3 million such smallholders in the country. Smallholders practice rain-fed agriculture, operate at low levels of productivity. Most smallholder production is committed to own-consumption, but there has been considerable growth in the marketing of both basic food crops and cash crops by smallholders.

The data gathered through the TIA is comprehensive. Small-and medium-scale farm surveys include data on household characteristics (household identification, and number of household members), access to services, associations, credits, and disasters effects, income indicators (salaried employment, self employment, and remittances and pension), production and sales of grains, and food security and household vulnerability.

For my purposes, I am primarily interested in the geographic location of each household. A pair of latitude and longitude coordinates was recorded for each household, which makes it possible to map out all respondents, and measure the distance between each pair by either Euclidean distance or nearest neighbor methods using ArcGIS spatial modeling software. In order to identify the role of spatial proximity in network relationships, it is necessary to have precise information on the geographic locations of adopting and non-adopting households.

I used data from the 2008-2011 Partial Panel Survey (PP2011), which is a partial survey of TIA 2007 and TIA 2011. My sample includes households interviewed in 2011 and a subset of households that were initially interviewed in 2007 and re-interviewed in 2011. The survey was conducted in the provinces of Nampula, Zambézia, Tete, Manica, and Sofala and includes data from 1,454 households. Table 3.1 shows number of households sampled in each province.

Table 3.1: The number of adopters in the sampled provinces

	Nampula	Zambezia	Tete	Manica	Sofala
Number	263	330	277	244	340
Percentage	18.1%	22.7%	19.1%	16.8%	23.4%

The data set includes information on farm identification, farm household characteristics, production and sales, access to services, information on price and production, and risk factors. Table 3.2 presents the variables included to study the adoption of improved maize varieties. In this sample, 66.6% of the household heads are males at an average of 44.35 years old with 2.98 years of education. Only 1.9% of them have had training about the new variety and only 27.9% of the sampled household heads have a paid job. A large proportion (65.5%) of the households have prior experiences with the new variety[15]. Adopters use either self-owned seeds (29.4%) or bought seeds (20.2%). Of all the households that used improved seeds, 29% sell their harvest and 48% use their harvest for own consumption. Regarding the sources of information, 15% of the households have received information from an extension service in the past 12 months. Price information comes from NGO (35.4%), radio (22.5%), extension (4.6%), and associations (4.5%). A small percentage (7.8%) of the sampled households are part of agricultural associations and 3.9% received agricultural credits from the government. Survey respondents reported three types of calamities that may have adversely affected maize production: drought, flood and cyclones, of which drought is the number one risk (34.4%), followed by cyclones (15%) and flood (6.7%).

In the next section, I discuss specific function forms for adoption, namely a spatial latent model with binary outcomes (or a spatial probit model). A spatial latent variable model is a specification where spatial correlation is introduced between the decision variables and/or in the error structure of the model (Anselin 2002). In terms of a formal model of variety adoption, a spatial regression model includes measures of others' variety adoption as an additional explanation variable, moderated by others' relationship with the farmer in question. Because the data are cross-section in nature, a spatial econometric model captures the importance of others in a single parameter.

3.4.2 Model

My econometric model is based on the assumption that social interactions can be modeled as spatial phenomenon. The analogy between social relationships and space is

[15] This sample consists of farmers who have adopted before.

not a new one, as Yang and Allenby (2003), Narayan, Rao, and Saunders (2011), and Richards, Hamilton and Allender (2014) consider explicitly spatial econometric models of social interaction. In my case, a spatial latent variable model is a formal representation of the equilibrium outcome of social and spatial interaction. Even though the actual dynamics of the interaction among agents (peer effects, neighborhood effects, spatial externalities) cannot be observed due to the single dimension of observations for a single point in time, the correlation structure that results in the equilibrium can be modeled (Brock and Durlauf 2001, Durlauf 2004).

The standard approach to modeling such phenomena is to develop a specification for an unobserved underlying latent dependent variable for each farmer, defined as ${y_i}^*$. The link between the latent variable and the observed discrete phenomenon (adoption) is obtained by specifying a threshold, *c*, such that y_i is observed whenever ${y_i}^* > c$. In general notation, a spatial lag model is written as (Anselin 2002):

$$ {y_i}^* = \rho \sum_{j \neq i}^{n} w_{ij} {y_j}^* + {x'}_i \beta + u_i \tag{3.20} $$

where ${y_i}^*$ is a $n \times 1$ vector of latent dependent variables, $\boldsymbol{W}$ denotes a spatial weight matrix that represents the spatial distance among farmers, $\boldsymbol{X}$ denotes explanatory variables that may contribution to adoption decisions as discussed in the previous section. In this expression, the latent variables are not the same as the discrete outcomes of adopting or not adopting, but an unobserved measure of utility that will lead to the observed adoption variable. In other words, it is the latent $\boldsymbol{W}y^*$ that is present in the actors' objective functions, but not the observed $\boldsymbol{W}y$. For example, in the context of my model, it is the unobserved profitability of the neighbors' parcels that enters in the utility function of a target input model, but not the observed input.

Let the threshold $c = 0$, and let y_i be the binary outcome whether a farmer will adopt or not, taking on the value of 1 (adopt) whenever $y_i^* > 0$. The threshold value is interpreted as the utility from adopting the new variety. Whenever the unobserved utility is greater than zero, a farmer makes the observed action of adopting. The probability of adoption is then:

$$Prob[y_i = 1] = Prob[\varepsilon_i < \rho \sum_{j \neq i}^{n} w_{ij} y_j^* + x'_i \beta] \tag{3.20}$$

While equation (3.20) shows the empirical adoption decision for a single farmer, the spatial dimension of the problem is best represented by showing all adoption decisions together, in matrix notation. In this way, the spatial relationships among farmers becomes clearer. In matrix notation, the latent spatial lag process is given by:

$$\boldsymbol{Y} = (1 - \rho \boldsymbol{W})^{-1} \boldsymbol{X} \beta + (1 - \rho \boldsymbol{W})^{-1} u \tag{3.21}$$

Where $\boldsymbol{Y}$ is a vector of binary outocmes of adotpion, $\boldsymbol{X}$ is a vector of explanatory variables, $\boldsymbol{W}$ is the weight matrix that indciates the distances among farmers, and u is an idiosycritic shock.

By definition, the reduced form is nonlinear in ρ and β and has a spatially correlated error structure (a spatial autoregressive structure). This means that that the value of *y* at any location *i* is not only determined by the values of *x* at *i*, but also of *x* at all other locations in the network (or in the case of neighborhood, all neighbors). As an illustration of the autoregressive effect, expand the inverse matrix term (for $|\rho| < 1$ and with a row-standardized $\boldsymbol{W}$), and using the expected value (since the errors all have mean zero) to arrive at:

$$E(\boldsymbol{Y}|\boldsymbol{X}) = \boldsymbol{X}\beta + \rho W\boldsymbol{X}\beta + \rho^2 W^2\boldsymbol{X}\beta + \cdots \rho^n W^n\boldsymbol{X}\beta \tag{3.22}$$

Where the expected adopting actions of farmers given the explanatory factors are mediated by the influence of nearby farmers, with closer neighbors having stronger effects and distant neighbor's having weaker effects.

The powers of ρ and $\boldsymbol{W}$ highlight the distance-decay effect because for higher orders of ρ, there is less influence placed on the spatial effect of others' behavior. This expression also shows how spatial effects can be both local and global. Local effects are represented by $\rho\boldsymbol{W}\beta$, which shows how the famer is influenced by his immediate neighbors. On the other hand, global effects are given by $\frac{\beta}{1-\rho}$, which shows how a farmer is influenced by the entire network. These differences highlight the power of using a spatial approach to study informational effects in a social network.

Clearly, the spatial weight matrix plays an important role in any spatial autoregressive model. There are many ways to define the individual elements of a spatial weight matrix, capturing the various ways in which neighbors, or other members of a social network, interact. Essentially, these weights formalize the neighborhood structure between the observations as a $n \times n$ matrix $\boldsymbol{W}$ in which each element $w_{ij}$is non-zero for the existence of a neighbor relation, and zero otherwise. By convention, the diagonal elements are set to zero, $w_{ii} = 0$.

The specification of the weight matrix is a matter of some arbitrariness and depends on both the context, and the objectives of the researcher (Anselin 2002). A range of suggestions have been offered in the literature, based on contiguity, distance, as well as more general metrics (Anselin 2006, Anselin and Rey 2006, Lesage 1998, Lesage 2010). Weights based on contiguity include rook and queen (from chess terms) continuities where only units that are adjacent to the focal unit in the rook/queen fashion are considered neighbors. A contiguity matrix is normally calculated from polygon data, that is, geographical units with boundaries. When two units share the same boundary they are considered neighbors. In my data, a pair of point GPS coordinates does not indicate boundaries. Therefore my contiguity matrix is calculated using the method of Thiessen polygons. A Thiessen polygon surrounding a given household is constructed by drawing lines between that household and all other households. These lines are then perpendicularly bisected in their middle and treated as "boundaries". Rook contiguity is then calculated with four neighbors that share boundaries directly with the concerned household[16]. The advantage of a contiguity matrix is that it addresses immediate neighbors, who are the most accessible information sources. Intuitively, direct neighbors should have a significant influence on adoption decision because of the convenience of learning from each other.

Distance-based weight normally refers to the Euclidean weight matrix, where the relationship between two units is measured by the relative distance between them. A Euclidean weight matrix is a truthful reflection of the geographical space, however, because the complexity of a relationship is not always dependent upon closeness, the Euclidean weight matrix may fail to represent the underlying travel expenses among the units of observation because it measures the direct distance from point A to point B. For

[16] This is actually the same as calculating the four nearest neighbors, which method I mention next.

example, consider a mountainous farming environment. Two neighbors may be only 1 mile distant as the crow flies (Euclidean distance), but the natural barrier may mean that a farmer 10 miles down the river is a much stronger influence. For this reason, I use an arc-distance matrix, which address not only direct natural distance from each other, but also slopes (or valleys) that provide extra distance (and not shown on Euclidean distance). For example, assume there are two households, each live on one side of a mountain. The Euclidean distance measures the horizontal distance between two houses, whereas the arc distance accounts for the distance to climb over the mountain. The advantage of distance-based matrix is that it represents the true travel distance between households. The significance of such a representation as the social proximity is questionable because nowadays, travelling does not bound farmers with respect to information exchange.

I derive a third matrix based on the concept of the “nearest-neighbor.” A nearest-neighbor matrix is one in which the nearest locations are considered as neighbors and coded as 1 to indicate they are likely to have a spatial relationship. An advantage of a nearest-neighbor weight is that, unlike continuity weights, a nearest neighbor matrix pertains to only immediate neighbors. This allows for investigation of relationships among households that transcends geographical limitations. Farmers can be spatially related by socio-economics rather than geographic distance. For example, the nearest neighbors of a focal farmer can be from the same geographic region, or they can be across regions, but possess some underlying economic similarities such as accepting loans from the same bank. For example, if accepting credits/loans is a crucial for farmers’ adoption decisions, farmers should express similar behavior based on the mutual loan policy.

I include three categories of explanatory variables into the ***X*** matrix, namely they are *farm identification* variables that indicates the located provinces, *household characteristics* variables that describes the socio-demographic of the households, *production and sales* variables that indicates the production history logistics, *non-social learning* that describes the sources of information besides social learning, and the *risk* variables that describes the risks for Mozambique farmers. The selected variables are presented in Table 3.2.

Table 3.2 Selected Variables and Descriptive Statistics.

Variables	Definition	Descriptive Statistics Frequency	
Farm Identification			
Nampula	The households belong to Nampula province		
Zambezia	The households belong to Zambezia province		
Tete	The households belong to Tete province		
Sofala	The households belong to Sofala province		
Household Characteristics	Gender of the household	1= male	66. 6%
HH_Gender		0= female	14.6%
HH_Age	Age of the household head		44.35 (13.856)
HH_Education	Years of education for the household head		2.98 (2.993)
HH_Training	The household had agricultural training in the past 3 months	1= yes 0 = no	1.9% 79.3%
HH_Job	The household head had salaried employment	1= yes 0 = no	27.9% 53.2%
Production and Sales			
Improve	Grew improved maize variety in 2011	1=yes 0= no	11.3% 65.6%
Impr_Before	Had grown improved maize prior to 2011	1= yes 0= no	11.5% 65.5%

ImprSeed_Own	Owned improved maize seeds	1= yes 0= no	29.4% 47.5%
ImprSeed_Buy	Bought improved maize seeds	1=yes 0=no	20.2% 56.7%
Impr_Sell	Sold the maize grown with improved seeds	1=yes 0=no	29% 48%
Non-social learning			
Info_Extension	Received information or advice from extension in the past 12 months	1=yes 0=no	15% 66.2%
InfoPrice_Radio	Price information from radio	1=yes 0=no	22.5% 58.7%
InfoPrice_Extension	Price information from extension	1=yes 0=no	4.6% 76.5%
InfoPrice_NGO	Price information from non-government organizations	1=yes 0=no	35.4% 45.7%
InfoPrice_Assc	Price information from agricultural association	1=yes 0=no	4.5% 76.7%
Mem_Assc	Member of agricultural association	1=yes 0=no	7.8% 73.4%
Credit	Received agricultural credits.	1=yes 0=no	3.9% 77.2%
Training	Received training in the past 3 months	1= yes 0= no	2.2% 97.8%
Risk Factors			
Risk_Flood	Lost crops due to flood in the past 12 months.	1=yes 0=no	6.7% 74.4%
Risk_Drought	Lost crops due to drought in the past 12 months.	1=yes 0=no	34.4% 47%
Risk_Cyclone	Lost crops due to cyclone in the past 12 months	1=yes 0=no	15.6% 65.5%

Among household identification variables, I expect the central regions to show a higher adoption rate as maize is more commonly grown the central regions. Socioeconomic attributes are variables that measure observed heterogeneity, and in that regard are likely to explain some of the observed variation in adoption rates. Within the set of socioeconomic attributes available in the TIA data, I expect education to be positively related to adoption because educated people have more access to information, and a better understanding of the market. If new varieties are indeed better, then higher education should correlate with adoption. Age of the household head has an indeterminate effect on adoption. While experience can play a very important role in that more experienced individuals may have more accurate prior about the new variety, making them more likely to adopt, older farmers can be set in their ways and less likely to switch to a new variety. Other exogenous variables include the availability of government information sources. I expect that receiving agricultural credits and being a member of an agricultural association have positive effects on one's adoption decision. Credits are given to farmers as an incentive to produce certain crops, and as a way to alleviate the damage caused by any natural disasters. Being a member of an association not only provides the opportunity to learn about new varieties, but also exposure to market information and advances in production methods.

Other farm-related attributes may also explain some inter-farm variation in new variety adoption. First, I included variables that represent whether a farmer's harvest was for own consumption or for sale. A farmer's own experience, independent of age, is also likely to be important. In that regard, I included a variable indicating whether the farmer had planted the variety in question in a prior year. This effect may be positive or negative, depending on the farmer's particular experience with the variety. If the experience was a good one, then the effect is likely to be positive, but negative if it was not. Risk may also slow adoption. In order to capture the influence of risk, I included the occurrences of drought, flood and cyclone in the past 12 months. In Mozambique, wildlife poses the most important risk, so I capture this effect by including a variable to indicate whether the crops were attacked by wildlife in the past 12 months.

Regardless of the definition of space, spatial econometric models cannot be estimated using conventional, ordinary least squares methods, due to the endogeneity inherent in their structure. In the next section, I describe an estimation strategy that exploits the unique nature of my data for identification.

3.4.3 Estimation

Spatial econometric models are powerful in that they are able to encapsulate a large amount of information in a relatively simple form, but at the cost of estimation complexity. For example, to including a spatially lagged dependent variable in a regression specification creates endogeneity because the choices made by everyone else in the network are correlated with unobservable factors, by definition. There are two types of spatial dependences: the lag dependence, which studies the influence between an individual (or unit) and his (its) neighbors, or the error dependence, which focuses on how unobserved observations are explained by the spatial relationships. In our case, information exchange is analog to the lag specification because information is transmitted among farmers and decay in the same process. A lag specification can capture the nature of such decay, or in other words, how is information reserved among neighbors, and then translate into the network influence of neighbors as a whole.

Modeling discrete choice in a spatial framework is challenging due to the violation of IID in the error distribution (because it is cofounded with a inverse matrix considering weight, or inverse Leontief matrix). Simulation methods are generally employed to get unbiased and consistent estimates asymptotically. For this study, I use the maximum likelihood estimation (MLE) method developed by Fleming (2000) and follow the procedure of Pinkse and Slade (1998) in that I first test the spatial dependency with my data. After rejecting the null of a traditional probit, I estimate a spatial probit (or a spatial latent model by Anselin (2010)). In this section, I introduce the estimation method first, followed by the specification test I employ to determine spatial lag dependency. Results are reported in the reverse order in order to establish the preferred form of the model prior to interpreting the associated parameter estimates.

MLE assumes the error term is normally distributed, or ε~N(0, Σ_{θ}). Using the general notation for equation (3.21) above, the log-likelihood function (LLF) for the spatial lag model is given by:

$$lnL = \sum_i [y_i ln\Phi(x_i'\beta) + (1 - y_i)\ln\big(1 - \Phi(x_i'\beta)\big)] \tag{3.25}$$

with $y_i = 0, 1$. Each observation on the discrete dependent variable y_i can be considered to be an independent draw from a binomial random variable with probability $\Phi(x_i{}'\beta)$. Since $1 - \Phi(x_i'\beta) = \Phi(-x_i'\beta)$, and using $q_i = 2y_i - 1$, this can be expressed as:

$$lnL = \sum_i ln\ \Phi(q_i x_i'\beta) \tag{3.26}$$

In the spatial case, the simple summation is no longer appropriate and the full multivariate density must be evaluated to obtain the log-likelihood. Consider *u* as the n×1 vector of multivariate normal random variables with variance- covariance matrix Σ. In order to generalize the censoring conditions for both values of y_i, set $\boldsymbol{Q}$ as a diagonal matrix with diagonal elements q_i defined above. The multivariate censoring conditions are:

$$u < Q(I - \rho W)^{-1}X\beta \tag{3.27}$$

These must be evaluated in a multivariate normal distribution with Σ respectively as $[(I - \lambda W)'(I - \lambda W)]^{-1}$. The corresponding log-likelihood can be expressed as:

$$lnL = ln\Phi_n[Q(I-\rho W)^{-1}X\beta; 0, \Sigma_\rho] \tag{3.28}$$

Where Φ_n denotes a n-dimensional multivariate normal cumulative distribution function with the upper bounds as the first term, mean 0, and variance-covariance matrix Σ.

To test for spatial lag dependence, I employ a Lagrange Multiplier (LM) test statistic for the spatial autoregressive process (Anselin 1988a). The LM test statistic is Chi-square distributed with one degree of freedom and is written:

$$LM_\rho = \{\frac{e'Wy}{\frac{e'e}{n}}\}^2/D \tag{3.29}$$

While e is the OLS residuals and the denominator D is:

$$D = \left[\frac{(\boldsymbol{WX}\hat{\beta})'[I-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'](\boldsymbol{WX}\hat{\beta})}{\hat{\sigma}^2}\right] + tr(\boldsymbol{W}'\boldsymbol{W} + \boldsymbol{WW}) \tag{3.30}$$

Where the estimates for $\hat{\beta}$ and $\hat{\sigma}^2$ are from OLS. The test statistic is asymptotically distributed as $\chi^2(1)$. For this test, the null hypothesis is H_0: $\rho = 0$, so the alternative is an OLS model. Failing to reject the null hypothesis indicates there is no significant spatial lag dependence, while rejecting the null hypothesis indicates the existence of a spatial lag.

The selection of weight matrix is arbitrary and subjective due to the underlying economic/geographic reasons, therefore current practice is in need of a formal criterion to select the spatial weight (Wheeler and Páez 2010). There are currently three different approaches for exogenously estimating the specification of weights: cross-validation (Brunsdon et al. 1996; Farber and Páez 2007), corrected Akaike Information Criterion (AIC, Fotheringham et al. 2002), and the specification LM test (Anselin 1988b) discussed above. Cross-validation (CV) is an iterative process that searches for the weight that minimizes the prediction error using a subset of the data or prediction. Because CV depends on the predicted value, it is compromised when the model has a poor prediction power (such as my case). On the other hand, AIC does not base on prediction of the response variable. It is instead based on minimizing the estimation error of the response variable. Unlike a LM test, AIC does not test against a null hypothesis. Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. In the selection of weight matrix, I use both AIC and the specification LM test. Results are presented in the next section.

3.4.4 Results

I first present a set of specification tests in order to establish the validity of my spatial model. I considered a number of alternative models, including a non-spatial probit model, a spatial model using arc distance weight (Model 1), a spatial model with rook contiguity (Model 2), and a model using nearest-neighbor weights (Model 3).

Comparing the LM Lag test value against the critical value suggests that the null is rejected with the specification of immediate neighbors and extended neighbors, but not in the specification of arc distance. Intuitively, it tests the residuals from a non-spatial probit model against a series of spatial models that differ only in their weight specifications[17]. The null hypothesis is that there is no spatial auto-regression among households' adoption decisions. Table 4 shows the results from comparing each specification against the non-spatial null model. A spatial model is indeed preferred in

[17] Mathematical expressions of the tests can be found in Anselin and Rey (2010).

explaining the sample adoption data using nearest neighbors (LM=5.956>${\chi^2}_1$) or rook neighbors (15.766>${\chi^2}_1$) as proximity among households. However, distance doesn't explain the adoption decision among households as the LM value is smaller than the critical value. This finding also supports *hypothesis 3* that there is likely significant spatial lag dependence among farmers' adoption decisions. Intuitively, when the null hypothesis is rejected in favor of a spatial lag, the spatial lag parameter should be statistically significant. The value of the spatial lag parameter represents the nature of spatial influence of specific weight. In this case, spatial lag parameters for both rook contiguity and nearest neighbors are statically significant. Table 5 shows the estimates results of spatial models using arc-distance, contiguity, and nearest neighbor definitions. Spatial lag parameters are significant for contiguity (*p* value of 0.0001) and nearest neighbor (*p* value of 0.0016), but not for arc-distance (*p* value of 0.5389). The magnitude of dependency is represented by the value of ρ, with nearest neighbor having a lower dependency (0.0625) and rook contiguity having a lower dependency (0.5086).

In term of goodness-of-fit, Model 2 posses the lowest AIC where as Model 1 has the highest. AIC supports that models using nearest neighbors and rook contiguity as weights perform better than the model that uses arc distance. To compare the goodness-of-fit of Model 2 and Model 3, I employ Log-likelihood Ratio (LR) to test the dominance among models. A LR test compares the performance of two models by taking the difference between the log-likelihood (LL) of two models and compares it to a chi-square distribution:

$$LR = -2(LL_{MNL} - LL_{RPL}) \sim \chi^2 \quad .$$

From Table 4 I conclude that Model 2 is the preferred model among the three models as the LR value between Model 1 and Model 2 is 6.392, which is greater than the critical value of 3.81, and the LR value (3.81) between Model 2 and Model 3 is also greater (or equal) than the critical value. Besides the LR test, I also employ the pseudo R^2. Lesage (1998) derives an expression for the coefficient of determination (R) for a spatial model that is analogous to the R^2 for OLS, but includes the spatial weight specification, so is referred to as a pseudo R^2. The pseudo R^2, like a traditional R^2, measures the portion of variation in data explained by the spatial model relative to the amount of total variation, and provides a measure of goodness fit. The pseudo R^2 indicates that rook contiguity model explains more of the variation (8.93%). Therefore I choose Model 2 as my preferred model for parameter interpretation.

Table 3.3 Diagnostics for spatial dependence test against a classic Probit

TEST	Arc Distance (Model1)	Rook Contiguity (Model 2)	Nearest Neighbors (Model 3)
LM Lag	0.377	15.766*	5.956*
Log-likelihood	-288.397	-282.005	-285.813
AIC	602.782	590.011	597.627
Pseudo R^2	7.34%	8.93%	8.47%

*significant at 0.05.

Theories of social learning imply the sign of the relationship among adoption decisions is ambiguous (Bandiera and Rasul 2006). However, I find in that network is always positively related to learning, regardless of whether the network is formed out of immediate neighbors or extended members. This is because both spatial lag parameters are positive (shown in Table 5), indicating that farmers tend to copy their neighbors' decisions (no matter what the decisions are). Therefore, having more adopters will encourage farmers to adopt while having more non-adopters will prevent farmers from adopting. This finding explains the clusters that are presented in Figure 1, where adopters tend to reside in close proximity, separate of non-adopters. Moreover, my results show that such clusters are not caused by regional difference, as none of the province is found to be significant in explaining adoption. Combining the two findings, I conclude that local networks play an important role in new maize variety adoption in Mozambique. To promote new varieties, the government needs to make sure that they indeed perform better that the traditional varieties. Once farmers receive positive feedback from the new varieties, they will spread out the words among their networks and other farmers will follow. In the case of my data, having previous experience with the new variety is negatively related to adoption, as evident by the negative sign of *adopted.* That is, on average, farmers who planted the new variety in 2007 are less likely (4%) to plan it again in 2011. Many reasons can contribute to the negative relationship between previous examination in 2007 and adoption in 2011: lower regulated price, bad weather condition or natural disasters. My two-period data does not allow me to conclude on this notion.

Table 3.4 Spatial models using different weights.

Variable	Arc-distance	Contiguity (rook)	Nearest Neighbor
Constant	0.0211	0.0211	0.0910
	(0.1399)	(0.6222)	(0.0270)
Education	0.0097*	0.0075	0.0009*
	(0.0038)	(0.0556)	(0.0038)
Training	0.3551*	0.3572*	0.3562*
	(0.0778)	(0.0000)	(0.0773)
Adopted	-0.0498	-0.0886*	-0.0535*
	(0.0323)	(0.0141)	(0.0321)
Extension	0.0682*	0.0710*	0.0699*
	(0.0299)	(0.0175)	(0.0298)
Nampula	-0.1142*	-0.0470	-0.1053*
	(0.0417)	(0.2882)	(0.0342)
Tete	0.0001	0.0074	0.0010
	(0.0304)	(0.8057)	(0.0302)
Zambezia	-0.0314	-0.0102	-0.0305

	(0.0307)	(0.7503)	(0.0303)
Association	0.0162	0.0084	0.0168
	(0.0387)	(0.8267)	(0.0385)
Credit	0.1102*	0.1360*	0.1140*
	(0.0561)	(0.0168)	(0.0558)
Risk_flood	0.0383*	0.0331	0.0422*
	(0.0436)	(0.4456)	(0.0433)
Sale	0.0432*	0.0389	0.0440*
	(0.0236)	(0.0971)	(0.0233)
ρ	0.5685	0. 5086*	0.0625*
	(0.9709)	(0.0163)	(0.0257)
Lagrange Multiplier	0.3774	15.7662*	5.9562*
Prob. of LM (H0: $\rho = 0$)	0.5389	0.0001	0.0014
Log-likelihood	-288.397	-282.005	-285.813
Pseudo R^2	7.34%	8.93%	8.47%

Note: an asterisk indicates significance at 0.05.

Non-social learning also plays an import role in explaining variation in adoption rates, which provides support for hypothesis 2. *Education*, *training* and *extension* are all found to be significant and positively related to adoption. *Extension* is positively associated with the households' adoption decision. More specifically, one year of education increases the adoption probability by 0.9%, having access to extension services increases the probability of adoption by 6%, and more importantly, having specific training about the new variety increases the probability of adoption by 20%. This also provides supports for hypothesis 1 that experienced farmers are more likely to adopt. That said, training about the new variety is significant, and positively related to adoption of new variety. Clearly, famers prefer direct, hands-on help with planting new varieties. Credit is found to be another factor that promotes adoption, meaning that government subsidies is crucial in convincing farmers to switch from a familiar variety to a new one. Moreover, I find the purpose of production to be important in predicting adoption.

Specifically, if a farmer produces for sale instead of own consumption, he is more likely to apply the improved variety. In Mozambique, small farms normally produce for own consumption where as medium and large farms produce for sale. The advantage of new maize variety is high yield, which provide an incentive for larger farms to adopt. On the other hand, larger farmers possess economy of scales, therefore are better at sharing the unexpected price fluctuation brought by the new variety.

Addressing the information barrier is one of the focuses in this chapter. My results indicate that farmers communicate with immediate neighbors and use their adoption decisions as reference. They also extend their network to include more households for the benefit of information exchange. Distance, however, does not impose a constraint on learning. In fact, farmers go beyond their immediate neighbors as to seek information in order to maximize their benefits from information exchange.

3.5 Conclusion

In this paper, I investigate social learning among a group of farmers who are assumed to be passive price takers with forward –looking behaviors. In a theoretical model, I show that farmers have strategic incentives to delay the adoption of a new variety in order to wait for more information from other farmers. In reality, however, farmers' network provides a positive learning effect. Having connections with other farmers provides positive information externalities, which encourages learning. However, because the new variety is different from the traditional variety that farmers are used to, training is crucial in convincing farmers to switch. In Mozambique, training (2.2%) and extension services (15%) are limited in availability, which have caused delay in adoption. Even worse, the reason those who have adopted before choose to switch back to the old variety can be due to inadequate assistance on the new variety.

I test some implications of my conceptual model using farm-level data from Mozambique. My model is explicitly spatial in that each farmer is a member of a social network, and is related to each other member in a social-spatial way. With this model, I find that non-social learning is important in adoption, especially training. Training is found to increase adoption by 20%. Also, educated farmers are more likely to adopt. Information, or the lack thereof, is critical in speeding the adoption of new varieties. Even though price information is broadcast through TV and radio, only about 2% of farmers have access to this information. Access to extension services and agricultural associations is also limited, which contributes to the information barrier.

Throughout, I focus the discussion around social learning as the underlying mechanism connecting decisions within a network. I do this because lack of information is a key barrier to adoption in my setting and because previous work has shown farmers learn from each other about the parameters of a new variety. My primary finding is that local networks act as important agents for information exchange. Farmers rely on their immediate neighbors for recommendations, and weigh the neighbors' opinions heavily. Farmers also go beyond immediate neighbors to build their networks for the benefit of information. Furthermore, I find that geographic distance is not a barrier to farmers, as distance does not play an important role in adoption.

My findings have a number of important policy implications. First of all, training has been found to be crucial in promoting the new variety. Compared to modern machinery and high-tech farming approaches in developed countries, direct, hands-on demonstration is a more viable approach due to the limitation in education and social media in rural Mozambique. Moreover, government credit is another way to encourage Mozambique farmers to switch to improved variety. The intuition is that, credits lower the cost of inputs for planting the new variety, and, therefore, lower the risk. Finally, although a majority of farmers produce for own consumption, those who produce for sale are more likely to adopt because of the high yield advantage of the new variety. In this regard, government should aid in the supply-demand of maize market and help absorb the end products from farmers.

While my analysis is tailored to the specific context of rural Mozambique, the findings have broader applicability. First, similar patterns of initial adoption decisions within and across social networks may occur in other economic environments in which a new technology is introduced, information is a key barrier to adoption, and individuals can be expected to learn about the new technology from others. In particular, the spatial nature of relationships between adopters should always be taken into account in contexts where social learning plays a key role.

This study is not without limitations. I focus only on maize adoption in order to keep the empirical exercise tractable in scope. Maize is a staple crop in Mozambique, but there are other cash crops such as cotton and cassava in which the adoption decision may be as or more important. Second, network costs are only implicit in my empirical model. A fully structural empirical model would endogenize the cost of forming and maintaining the network. In the context of a spatial probit model, however, endogenizing costs in this way is similarly intractable. Finally, my spatial model is based on the assumption that local networks are constituted by nearest neighbors. Other network structures are possible, and should be considered in future work.

CHAPTER 4. ESSAY3: MODELING PRODUCT CHOICES IN A PEER NETWORK

4.1. Introduction

Consumers make decisions in a world of uncertainty with imperfect information and only partially formed preferences for attributes they know little about. They search for information as a means of reducing uncertainty and improving the likelihood that they will be satisfied with their purchases. Information is provided by advertising, physical media, online media, and, perhaps most importantly, from peers. Research in a range of fields has shown that peers are critical in shaping preferences and choices (Manski 2000; Sacerdote 2001; Zimmerman 2003; De Giorgi, Pellizzari, and Redadelli 2010; Kuhn et. al. 2010; Richards, Hamilton and Allender 2014). Indeed, peers have been shown to be important in the apparent clustering of obesity (Christakis and Fowler 2007; Cohen-Cole and Fletcher 2008; Trogdon, Nonnemaker, and Pais 2008), the popularity of otherwise-unheralded movies (Reinstein and Snyder 2005; Moretti 2011), retirement plan participation (Duflo and Saez 2002, 2003), health-plan choice (Sorensen 2006), investing in the stock market (Hong, Kubik and Stein 2004), performance in college (Sacerdote 2001; Zimmerman 2003), behavior in school (Evans, Oates and Schwab 1992; Soetevent and Kooreman 2006), or new product purchases (Mayzlin 2006; Godes and Mayzlin 2004, 2009). In each case, the exact mechanism through which peers exert influence on others differs. In this chapter, I compare two mechanisms through which peer networks may operate, namely the strength of social ties and perceived peer expertise to draw implications regarding how consumers' preferences are revised through peer recommendations.

My aim is to identify true peer (endogenous) effects that cause individuals to have similar preferences. Importantly, I disentangle endogenous effects from contextual and correlated effects (Manski 1993) that are often overlooked in studies of social influence. I do so by conducting a randomized two-stage experiment to elicit subjects' willingness to pay (WTP) for activity trackers. I analyze preference changes between the first and second stages using a spatial econometric approach that helps identify the peer effects I seek. By controlling for correlated and contextual effects, my experimental approach is able to cleanly identify significant endogenous effects.

Clear identification is necessary in order to isolate the precise mechanism through which peer effects operate. In my experiment, I collect data on source expertise, i.e., how subjects evaluate the apparent expertise of others, and tie strength, i.e., how well each subject knows the others. Afterwards, I estimate the importance of each in shaping endogenous peer effects.

I find that perceived source expertise is important in influencing preferences while the closeness of social relationships is not. These results are consistent with the findings of Freeman (1957) and Pornpitakpan (2004) in that sources having more credibility dimensions are more influential than those having less credibility on readership scores[18]. Moreover, in contrast to the findings of Richards, Hamilton and Allender (2014), people who are close in social space are not likely to have a significant influence on each other. This suggests that subjects who are perceived as credible are not

[18] A source high in expertise, as compared to one low in expertise, appears to lead to positive attitudes toward the endorser and the advertisement (Braunsberger, 1996). Degree of perceived credibility of the source influenced recipients' intention to use suggestions made by the source as to how to improve performance (Bannister, 1986) and the acceptance or rejection of the suggestions from the source (Suzuki, 1978).

necessarily close friends, confirming that weak ties can be more important in transferring information than strong ties (Granovetter 1973, Levin and Cross 2004).

I contribute to the literature on social learning and new product introduction in a number of ways. First, I contribute to the theoretical marketing literature in that I find social learning via source expertise to be a more effective form of peer recommendation than tie strength. Granovetter (1973) theorized that weak ties are more likely than strong ties to be bridges to socially distant regions of a network and, therefore, new information. I provide further insight into the way in which communication networks emerge. More specifically, the decision to seek information from a specific other is informed by characteristics of the relationship between the seeker and her peers based on the perceived credibility. Second, my findings are likely to be of interest to marketing practitioners because new-product introduction is a fundamental marketing task. Word-of-mouth and viral reputations are among the most effective (Ferguson 2008, Bruyn and Lilien 2008). My findings suggest that face-to-face peer influences, where peers are regarded as experts, are significant in revising others' preferences. Successful practices such as Yelp and TripAdvisor are examples that utilize source expertise to promote certain restaurants and places of interests. I also contribute to the methodological literature on estimating social learning effects as I introduce a new method of estimating peer effects using spatial econometrics. While some authors have used the concept of space to help identify peer effects (Yang and Allenby 2003; Richards, Hamilton, and Allender 2014) I extend the definition of spatial relationships to consider tie strength and source expertise to measure peer influences. In this regard, I highlight an essential point of the social learning literature, namely that how relationships are defined is essential to understanding the nature and power of social influence within a network.

In the next section, I frame my contribution in terms of the existing literature on social networking and econometric methods of identifying peer effects. Building on previous studies, I then describe the experiment, and provide some summary evidence of the power of peers to shape behavior. In the following section I present my econometric model, and show how it is able to identify peer effects separately from contextual and correlated effects. The fourth section presents the empirical results, and discusses some of the implications for marketing practice. A final section concludes and offers some suggestions for further research in this area.

4.2 Background

The context for my investigation concerns the effect of peer influences on marketing an innovative new technology product, activity trackers in my case. Peer effects in a marketing environment are commonly referred to as word-of-mouth (WOM), and operate through mechanisms that include source expertise (Bansal and Voyer, 2000 and Gilly et al., 1998), tie strength (Granovetter 1973, Brown and Reingen, 1987 and Frenzen and Nakamoto, 1993), demographic similarity (Brown and Reingen, 1987), and perceptual affinity (Gilly et al., 1998). I focus on the first two of these mechanisms: source expertise and tie strength.

Source expertise refers to the credibility or believability of a particular source of information. For example, recommendations provided by professionals and authorities are often perceived as more reliable. **Tie strength** refers to the closeness of the relationship between the individuals exchanging WOM. Logically, consumers are more likely to follow recommendations from people who they know and trust. However, Granovetter (1973) finds that weak ties between groups provide the greatest increment in information. Intuitively, weak ties can have a stronger effect because individuals tend to have weak ties with people from backgrounds that differ from their own and are, therefore, more likely to provide new information. On the other hand, people with strong ties are likely to share a similar background, so are less likely to introduce new information.

Empirically and conceptually, the mechanisms through which WOM operates are difficult to disentangle because of the "reflection problem" (Manski, 1993). The reflection problem "…is similar to an inferential problem that occurs when one observes the almost simultaneous movements of a person and of his image in a mirror" (Manski 2005, p. 2). When observing only an individual's movements, you cannot tell whether the mirror image is causing the person's movements, if the mirror is reflecting the person's movements, or if they are happening together. By the same token, when observing similar behaviors in a group, the analyst cannot always tell whether phenomena are caused by individual heterogeneity in preference, true peer effect, demographic similarities (contextual effect), or other effects.

Peer effects arise when "an agent's preference ordering over the alternatives in a choice set depends on the actions chosen by other agents" (Manski 2000). However, individuals tend to behave similarly to those around them for three reasons: Either because the individuals come from similar backgrounds (contextual effects, or demographic similarities), experience the same unobserved environmental influences (correlated effects), or are truly influenced by peer behaviors (endogenous effects)

(Manski 1993). Endogenous, contextual, and correlated effects represent fundamentally different behavioral pathways, each with different implications for changing behavior (Manski, 1993; Brock and Durlauf, 2001; Mofitt, 2001; Soetevent, 2006). Consider a student's college GPA as an example. There are endogenous effects if, all else being equal, an individual's GPA tends to vary with the GPA of the students he or she associates with. Endogenous effects arise either through learning, modeling others behaviors, mimicry, or some way in which an individual's behavior changes as a direct result of observing others. On the other hand, contextual effects arise if the student's GPA tends to vary with the socioeconomic background of the reference group. For example, if all lower-income students in a district have lower GPAs than higher-income students, then the effect measured by group-association is likely due to contextual effects. Third, correlation effects arise if students in the same school tend to achieve similar GPAs because they use the same after-class tutoring system. Although the students may be from different socioeconomic classes, and observe different peer behaviors, if they all benefit from the same exogenous influence, then this is a correlated effect. Distinguishing between these three types of effects is important because only endogenous effects allow for one individual to influence the outcomes of others in a network. Indeed, the term "peer effects" only properly refers to the presence of endogenous effects.

Attempts to separate peer effects from contextual effects and correlated effects usually involve experiments involving random assignment to different social groups (Duflo and Saez 2002, 2003; Kuhn et al 2010, 2011), econometric modeling (Yang and Allenby 2003; Lee 2007b; Lin 2010), or a combination of the two (Narayan, Rao and Saunders 2011; Richards, Hamilton and Allender 2014). For example, Duflo and Saez

(2003) study the impact of peer influence on the choice of retirement plans. In their field experiment, staff members at a large university were randomly assigned to work in different departments. Because participation rates in different retirement plans tended to cluster by department, their hypothesis was that discussion among staff members was primarily responsible for the observed clustering effect. The random assignment of staff members allowed the authors to identify peer effects, given that the participation rates of those in treated groups were twice as high as the participation rates of subjects in non-treated groups. In a similar randomization exercise, Kuhn et al (2011) randomized only the treatment group in a Dutch income-shock experiment and were also able to effectively identify peer effects. Randomization avoids the reflection problem by ensuring that common unobserved factors that could have caused similar behavioral outcomes were indeed random, meaning that the observed changes in behavior could only be due to the treatment effect. Randomized field experiments are one way to ensure peer effects are identified, but what appears to be random may instead be caused by an unobserved, endogenous factor.

Lab experiments provide researchers with more control over attributes of the sample and the environment in which decisions are made. For example, in a context similar to the one that frames our analysis, Narayan, Rao and Saunders (2011) use a two-stage conjoint choice experiment in which they measure subjects' willingness to pay for electronic book reader attributes. They collected subjects' initial preferences in the first stage, then asked subjects to identify their influencers. In the second stage, which took place two weeks later, subjects were shown the choices of their self-reported influencers and asked to choose from the same choice sets again bearing in mind the choices of influencers. They found that peer influence caused subjects to significantly revise their valuation of several attributes, and that this influence grew with the number of peers. One limitation to their experiment is that the waiting period between first and second stage was weeks, during which subjects' preferences could have changed for other reasons than the peer effect, or they could have forgotten why their preferences were initially formed as they were. Such lab experiments are advantageous because they allow researchers to select a sample that ensures that peer effects are identified. However, there is a limit to the ability of randomized experiments to eliminate confounding factors that appear as peer effects. Often, a common environment (that causes contextual effect) or unobserved factors – factors that can be subjective – can confound identification.

In order to address the problem of unobserved confounding factors, appropriate econometric models can help identify peer effects within experimental environments in which random assignment may be questionable. In general, econometric models of social interaction are of two classes: Bayesian updating (Yang and Allenby 2003; Narayan, Rao and Saunders 2011) and spatial techniques (Anselin 2002; Lee 2004; Lee 2007; Lin 2010). Bayesian updating assumes that agents observe signals from others before making their decisions. In other words, even when an individual has a prior, personal preference, she is very likely to update her preference based on the observations of her peers. Therefore, a Bayesian mechanism is an intuitive way to study how referent individuals in the social network assume leadership in promoting new products and measuring how followers copy their choices. In reality, however, assuming there is an order of influence can be problematic because peer influences often happen simultaneously and are reversible among members without one or even a few influential people leading the trend.

In this chapter, therefore, I apply a spatial model that allows for simultaneous influence among members of a social network. The influence is simultaneous in the sense that peers impact each other's choices.

Spatial models are more appropriate for simultaneous interactions (Richards, Hamilton, and Allender 2014) because they embody the type of feedback effects that are typical of social systems. Spatial regression models are designed to accommodate interdependence between spatial units (individuals, in our case) with cross-sectional data. Interdependence arises when the value of the dependent variable corresponding to each cross-sectional unit is assumed, in part, to depend on a weighted average of that dependent variable corresponding to neighboring cross-sectional units (Kelejian 1998). This weighted average is a spatial lag of the dependent variable, analogous to a temporal lag in a more usual time series model. When the root of this spatial lag is allowed to be an unobserved, estimated parameter, the result is a spatially autoregressive (SAR) specification. Because of the spatial lag term, simultaneous interactions are easily incorporated into the model, making it appropriate to study peer effects.

Spatial models are able to provide important insights with a minimum of parameterization. Slade (2004) provides an intuitive explanation of what the spatial lag effects are in this type of model. A spatial lag model is a formal representation of the equilibrium outcome of process of social space. The inclusion of the spatial lag is similar to an autoregressive term in a time series context in that, similar to the way in which a time lag is a weighted average of previous time periods, a spatial lag is a weighted average of neighbor characteristics. Unlike time dependence, spatial dependence is multidirectional, indicating feedback effect and simultaneity (Slade 2005, Anselin 2009).

When a person is spatially dependent on her peers then her decision is determined by her own characteristics and the decisions of her peers. The magnitude of such spatial dependence can be measured by the spatial lag parameter, which is bound between zero and one (in absolute value). A spatial parameter of one means that she is completely dependent on her peers while a spatial parameter of zero means she is completely independent of her peers. A positive spatial lag parameter indicates a person is positively influenced by her peers' decisions. By the same token, a negative spatial lag parameter means that a person tends to choose in ways that are opposite of her peers.

The classic SAR model departs from Manski (1993) by measuring peer variables as the weighted averages of observed peer outcomes and characteristics instead of group expectations. Group expectations create linearity in Manski's (1993) model, which further exacerbate the "reflection problem." Including weighted averages as explanatory variables is important because weighted averages introduce non-linearity, and non-linearity facilitates identification of endogenous and contextual effects (Bramoulle et al. 2009). Lee (2007b) demonstrates that both endogenous and contextual effects are identifiable in this model, but only when group sizes are not constant. Non-constant group sizes are necessary in this model, because it assumes each member of the group is the same "distance," in a social sense, from each other member. In this way, Lee's (2007b) model assumes that each group member is equally influenced by her peers, which is not realistic in practice. On the other hand, Lin (2010) addresses this somewhat unrealistic assumption by considering an arbitrary network structure based on observed relationships -- networks that are asymmetric through non-reciprocal peer nominations. For example, people who you consider as close friends may not necessarily have the same rating of relationship for you, which creates an asymmetrical and non-reciprocal social weight matrix. Heterogeneity in such arbitrary networks provides sufficient variation to identify peer effects separate from contextual effects (Bramoulle et al. 2009). Lin's (2010) model, however, confines interactions only within the group and ignores the out-of-group interaction.

In this chapter, I use individual-specific willingness to pay values calculated from data gathered in a choice experiment to estimate true peer effects. I use a spatial autoregressive (SAR) model similar to Lee et al. (2010) and Lin (2014), which includes endogenous effects, contextual effects, and group fixed effects, as well as a spatial autoregressive error process. Because subjects in my experiment are randomly assigned, and interpersonal relationships vary among individual subjects, peer effects are well identified in my model. The group fixed effects term, along with the spatial autoregressive disturbances, captures possible correlated effects caused by confounding factors. Furthermore, I investigate possible heterogeneity among peers, which is an important issue related to the specification of the spatial weights matrix that summarizes the network structure. To account for possible heterogeneity among peers, I consider different specifications of the weighting matrix, namely tie strength and source expertise.

4.3 Conceptual Model of Endogenous Peer Effects

The importance of peer effects has been examined empirically in the context of educational behavior (Sacerdote, 2000; Kremer and Levy, 2003; Hanushek et al., 2003), crime (Sah, 1991; Glaeser et al., 1996), teenage pregnancy (Evans et al., 1992; Crane, 1991; Anderson, 1991), and purchase behaviors (Narayan, Rao and Sanders 2011; Bollinger and Gillingham 2012; Richards, Hamilton and Allender 2014). Marketers often attempt to account for peer recommendations when determining the targeting and intensity of marketing activities (Bollinger and Gillingham 2012). The idea is that the utility an individual receives from pursuing a given activity depends on the recommendations of the other individuals in the person's reference or peer group (Manski, 1993; Becker 1996; Brock and Durlauf 2001; Glaeser and Scheinkman 2001).

Individuals are indeed strongly motivated to reduce the amount of effort they exert during the decision-making process in information intensive environments. Therefore their behavior may be directly influenced by effort-saving, easily available cues such as peer recommendations (Smith, Menon, and Sivakumar 2005). Prior research suggests that consumer choices arise from various information cues at the time of decision making, such as information via Web site background (Mandel and Johnson 2002) or attributes that just happened to be included in the recommendation of an electronic agent (Häubl and Murray 2003). Significant preference revision in the treatment groups will imply that subjects indeed look for easy cues during peer recommendation to revise their preferences. Therefore, I propose that positive recommendations by members of a social network is likely to influence another members' behavior in the same direction.

H_1: Peer recommendations will lead to preference revision for the recommended option if individuals regard the attribute in question to be salient to the choice decision.

The decision to seek information from someone when facing a new problem or opportunity is likely affected by the closeness between the seeker and her peers. Research on homophily indicates that people are more likely to have social ties (especially strong ones) with those similar to themselves (Marsden 1990, Zenger and Lawrence 1989, Brass

1995). Social network researchers have examined the role of weak versus strong ties in the acquisition of novel information. Granovetter (1973) argues that weak ties are more likely than strong ties to be bridges to socially distant regions of a network. Subsequent research on the importance of weak ties has demonstrated that they can be instrumental to finding a job (Granovetter 1973; Lin 1982, 1988), individual advancement (Burt 1992, 1997, 2000), and diffusion of ideas (Granovetter 1982, Rogers 1995). More recently, however, attention has shifted to the role of strong ties (Krackhardt 1992). Hansen (1999), for example, has demonstrated the importance of strong ties in transferring tacit, complex knowledge across departmental boundaries in an organization.

Tie strength should have a substantial effect regarding the influence of WOM communications (Brown, Broderick, and Lee 2007; De Bruyn and Lilien 2008). Strong ties are more likely to transfer useful knowledge (Levin and Cross 2004) and thus have more influence on others than do weak ties (De Bruyn and Lilien 2008; Smith, Menon, and Sivakumar 2005). Therefore, it is reasonable to expect that a strong tie between an individual and his or her peers is more likely to lead to preference revision than is a weak tie. I follow previous literature (Freeman 1987, Jackson 2008, Richards, Hamilton and Allender 2014) by defining tie strength in terms of the social distance (relationship) between two subjects. The relationship may be very close, for example, between relatives, or very casual, such as with acquaintances or strangers. The second hypothesis that follows from this theory is that:

H_2: Preference revisions will be greater for strong ties relative to weak ties.

It is also important that the information seeker is influenced by her perception of another person's credibility when making a decision about novel products. Knowledge of another person's expertise is a standard variable in the transactive memory literature, which identifies knowing where information is stored as a basic requirement of performance in distributed knowledge systems (Borgatti and Cross 2006). A consumer's subjective feeling of being influenced by the recommender may depend on how she feels about the recommender, or the perceived source credibility (Smith, Menon, Sivakumar 2005). Highly credible sources usually lead to more behavioral compliance than low-credibility sources (Pornpitakpan 2004). The dimensions of source credibility consist of expertise, which refers to the extent to which a speaker is perceived to be capable of making correct assertions (Hovland, Janis, and Kelley, 1953). Besides expertise, trustworthiness, defined as the degree to which an audience perceives the assertions made by a communicator to be true, is another important antecedent of behavior that demonstrate credibility (McKnight, Choudhury, and Kacmar, 2002).

I investigate how perceived credibility in the recommender impacts consumer preferences of activity trackers. Generally, those receiving information from opinion leaders associate the correctness of the information with their perceptions of the opinion leader's expertise in that particular domain (Feick and Higie, 1992). Therefore, I expect that perceived credibility, will positively influence a subject's preference revision, so that:

H_3: Perceived credibility is positively related to revisions in attribute preferences.

To test these hypotheses, I collect social network data from individuals in treatment groups and in control groups. I use a two-stage randomized experiment to introduce peer recommendations in the treatment groups but not the control groups. Through variation in the nature of the ties among subjects, the level of credibility perceived to be associated with each subject, and the levels of expertise and credibility in the recommender, I am able to analyze how social relationships influence preferences for product attributes.

4.4. Experimental Procedure

I conduct a social choice-based conjoint (CBC) experiment in order to examine the effect of peers on activity-tracker attribute valuation. A CBC experiment is ideally suited for my purposes because it permits an examination of tradeoffs among products that contain multiple attributes. Activity trackers not only contain multiple attributes, but are inherently complex, new to the market, and relatively little is known about them.

I use a CBC approach for a number of reasons. First, choices are a more realistic representation of true preferences compared to the ratings data gathered through other experimental methods (Rao 2007). In reality, consumers choose among different combinations of attributes and levels within each attribute, while they will only rarely rank their preferences when making a purchase decision. Furthermore, models estimated with choice data allow researchers to predict choice shares, which are of more interest to marketers. More importantly, choice models estimated in random utility form allow the derivation of willingness to pay (WTP) for product attributes.

The CBC experiment involves two-stages: The first stage is used to establish baseline attribute preferences for each subject, and the second measures the revision in these preferences based on exposure to input from others in a lab setting. A two-stage design is necessary, because eliciting peer effects requires an experimental mechanism wherein preferences expressed by a subject are allowed to be influenced by exposure to the choices made by others. I follow Narayan, Rao, and Saunders (2011) and Richards, Hamilton, and Allender (2014) in adopting this two-stage framework.

In the first stage, I elicit preferences for each attribute. Namely, I ask subjects to imagine they are in a store planning to buy an activity tracker. Each subject is presented with 12 choice sets, each with 4 alternatives, and an additional “no buy” option. Subjects are then asked to choose the preferred alternative or to choose “none of these.” After the first stage, I allow subjects in the treatment groups to express their preferences to others in the group, while subjects in the control groups are not allowed to communicate. After the exchange of information, subjects are asked to make the same choices again (stage two).

In general, a two-stage design addresses some of the most important challenges presented in the elicitation of peer effects. First, I choose subjects based on their major, which is assumed to be unrelated to their preferences for activity trackers. In this way, I control for the endogeneity of group formation. Second, I collect information on subjects’ social background and social proximity in order to control for unobserved correlation effects. Background information helps to filter out the contextual effect, addressing some of the unobserved factors. Third, measuring respondents’ preferences in both stages provides an opportunity to clearly identify endogenous peer effects, independent from any other influence that may have caused the observed preference revisions between the two stages.

Finally, I address the problem of simultaneity by estimating the extent of preference revision using a spatial-econometric model of choice utility. At the core of any spatial econometric model is a spatial weight matrix that captures the degree of relationship among the units of observation. In my case, the spatial weight matrix is a social-weight matrix as it measures the nature of the social relationship among sample members. This social weight matrix accounts for the social proximity of all network members. This approach also accounts for the simultaneity of sample members' decisions. Simultaneity arises in models of social influence because each subject influences, and is influenced by, all other subjects. Simultaneity is fundamentally an econometric problem as the choices made by others is both an explanatory variable, and an endogenous one (Anselin 2009). A spatial econometric approach is able to address this problem as it measures the degree of association between subjects, but excludes self-influence. My data captures the type of purely spatial influences that are easily identified using the spatial econometric approach developed in the next section.

In order to identify peer effects, the product in question must have attributes that sharply differentiate variants, its attributes must be relatively complex, and it should be somewhat new to the market so it is not well understood. In my experiment, subjects choose among activity trackers. Activity trackers represent an ideal product for studying the influence of others on consumer choices because they are indeed relatively new, innovative, and not well understood. Moreover, they are highly differentiated, and this differentiation rests on a small set of important attributes. Activity trackers combine modern technology with consumers' health awareness, so they are at the nexus of two trends – technology and healthy lifestyle – that should assure subjects' inherent interest in the products themselves. Trackers work on their own, or with handheld devices to provide biometric feedback, such as calories burned, steps, and miles walked or run. In addition, they show the users' daily activity goals so they can monitor how much activity is required in order to reach their goals. In this regard, activity trackers help users to become more aware of their health status.

Based on prior research, the salient attributes of activity trackers are price, style, brand, and function (Oh 2014). I chose four major brands of activity trackers on the basis of popularity: Nike, Fitbit, Jawbone, and Garmin. According to a report released in 2013, Nike, Fitbit and Jawbone accounted for 97% of all activity tracker sales (Mobilehealthnews.com). A Google shopping search revealed that reasonable price points include $49.99, $99.99, $129.99, and $199.99. Priced at $49.99, the Jawbone Up Clip on tracker attracts price-sensitive consumers. This type of tracker provides the basic function of tracking calories, but is limited by its design because the clip-on is not particularly well suited to intense exercise such as running. With slim wristband designs, Fitbit, Jawbone, Nike and Garmin all have trackers that are priced at $99.99 and $129.99. Newer versions are introduced every year, so older trackers are priced below the new models. Also, trackers that add emerging functions, such as sleep pattern tracking, are usually priced slightly higher. Trackers that are priced at $199.99 and above are often equipped with superior functions. For example, Fitbit Surge, priced from $200 to $249 on Google shopping, is built for multiple sports with a watch-type full OLED screen that will display calls, texts and notifications. Garmin Vivoactive, priced at $249.99, has GPS to accurately track running, cycling and swimming with live pace and distances. Trackers also vary in style, from watch-type, to wristband, and clip-on. Functions vary as the market has not yet settled on the core purpose of activity trackers.

I include both basic and innovative functions, such as the ability to record calories, the ability to record sleep, track running routes via GPS, and the ability to receive texts and emails. With these attributes included, my experiment spans both, the set of attributes that are relatively common and well-understood to those that are newer and likely to be not known widely. Each of these attributes, and levels, are shown in Table 4.1.

Table 4.1 Attributes and levels.

Attributes	Levels
Brand	Fitbit
	Jawbone
	Nike
	Garmin
Design	Clip-on
	Wristband
	Watch
Function	*Recording basic calories*
	Recording calories and sleep patterns
	Recording calories and text/email messages
	Recording calories and GPS locations
Price	$49.99
	$99.99
	$129.99
	$199.99

Clearly, the number of attributes and the levels of each imply more attribute combinations than can be described in a comprehensive way. In fact, a series of attribute profiles using a full factorial design that includes all combinations of the attribute levels consists of 196 (4*4*4*3) profiles to be evaluated by each respondent. A full factorial design is able to estimate both the main effects and interaction terms in the utility function, but 196 combinations is too many to present to the participants. Therefore, I use a fractional factorial design in which subjects were presented a subset of 48 combinations. My design is fully orthogonal in that it allows for the estimation of all main effects included in the study. The design is blocked in four blocks, so that each individual receives a balanced subset of profiles, namely 12 choice sets. I used SAS OPTEX to generate an orthogonal factional factorial design of 48 with a D-efficiency score of 80.64%. An example of one of the cards is presented in Table 4.2.

Table 4.2 Example Choice Set

	A	B	C	D	E
Brand	Garmin	Jawbone	Nike	Fitbit	None of
Design	Clip-on	Watch	Wristband	Clip-on	These
Function	Cal+Sleep	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$129.99	$49.99	
Choice					

One month before the experiment, a survey was sent out using Qualtrics to all students in the W.P. Carey School of Business and Ira A. Fulton School of Engineering at Arizona State University. The survey included a brief introduction to the experiment, detailed instructions, and an estimate of the time that would be required to complete the experiment. With the Qualtrics[19] online sign up system, I collected basic information from respondents who were willing to participate.

Respondents were required to be over the age of 18, ASU students, and able to communicate in English. The online sign-up procedure generated 80 eligible respondents. In order to keep the groups of a manageable size, I selected 20 people for each group. Before assigning groups, I asked each subject if they would be coming with someone they knew, and assigned people who knew each other to the same group. The group assignments resulted in four groups in total. I randomly selected two groups to be

[19] See appendix C2 for the Qualtrics recruiting scripts.

treatment groups and the other two served as control groups. Among the 80 responses I received, I then randomly selected 40 to be assigned to two control groups and another 40 to two treatment groups.

Randomization is introduced in the selection process in order to identify pure peer effects, or endogenous effects in Manski's (1993) terminology. Random assignment ensures that the difference in responses between the treatment and the control groups is due to the treatment alone, and not some pre-existing conditions that could have caused behaviors among group members to be similar (Duflo, Glennerster and Kremer 2007). For example, if I included mostly physical education majors in the treatment groups, and non-physical education majors in the control groups, and find significantly higher willingness to pay for activity trackers for the treatment groups, then this result could be attributed to the fact that physical education majors have more health awareness, and not due to pure peer effects. Subjects in the experiment are randomly assigned because students in the same faculty and year are likely to know each other, but the fact that they are classmates is unrelated to their preference for the particular item under study. Like Duflo and Saez (2002, 2003), sample is chosen based on their departments and major choices are not made regarding consumer products. In this sense, social relationships among students in a particular major represent a random assignment with respect to consumer-product attributes.

In the treatment groups, subjects were exposed to the attribute preferences of others between stages 1 and 2 of the experiment, whereas subjects in the control groups received no input before their stage two choices. Hence, the treatment effect measures the extent of peer influence, relative to the control in which no peer influence is allowed.

In each treatment group, I asked the subjects to talk about their choices, and why they made them. In addition, I asked the subjects to discuss the factors that influenced their choice of attribute packages. For example: what is the first thing you look at when you buy a tracker, the style or a specific function? The discussion regarding attributes was purposefully robust, with more experienced subjects often sharing firmly-held beliefs regarding the superiority of trackers they preferred.

Following the discussion, subjects were asked in stage two to again make their preferred choices between alternatives from the same choice sets as in stage one. Subjects in the control groups were not allowed to discuss their choices, but were instead asked to read an article on an unrelated topic. Diverting their attention from the task at hand was intended to take subjects' mind off the choices from stage one. After reading the article subjects also made their stage two choices. Both the peer discussions for the treatment groups and the reading for the control groups took 10 minutes. The entire experiment took approximately 35-40 minutes. After the choice experiment in stage two, I collected socio-economic and demographic data that is used to control for unobserved heterogeneity in the econometric choice model described below.

In any social experiment, characterizing relationships among the subjects forms a critical component of the analysis. These relationships form elements of the social relationship matrix. In a spatial model, these measures form the social "weights" that are used to filter out contextual effects, and to identify peer effects. I gathered data measuring closeness and source-credibility. Variation in "closeness" identifies tie strength because people who are closer to each other are characterized by stronger ties. More specifically, I measure closeness by asking subjects to report how well they know each of the other subjects. I follow the relationship measure of Richards, Hamilton and Allender (2014) where tie strength is defined as how well the subjects know each other, rated on a 5-point scale ranging from "Do not Know" (tie strength = 1) to "Know Very Well" (tie strength = 5)[20]. By choosing subjects that have the same major, the experiment is likely to include a range of relationships, from emergent "best friends" to only casual relationships. Variation in "perceived credibility" identifies source credibility (perceived expertise and trust) because people who are perceived as credible information carriers serve as opinion leaders, whose opinions are thought as more important. I measure source credibility by asking subjects to report how reliable they think each of the other subjects is. Reliability is measured on a 5-point scale from "Not Reliable" (reliability =1), "Somewhat reliable" (reliability =2), "indifferent" (reliability =3), "Somewhat reliable" (reliability =4), and "Very reliable" (reliability=5) (Bannister 1986, Borgatti and Cross 2006). This provides me with an assessment of source credibility from "most credible" to "not credible".

Among the 80 invitations sent out, 63 subjects completed the experiment in a useable way, leading to a turnout rate of 78.75%. Each subject provided 120 observations, resulting in a total of 7,560 observations. I follow Hensher, Rose and Greene (2005) in determining whether N=63 is an acceptable sample size. The minimum threshold for an acceptable sample size is defined as *n*, which is determined by the desired level of accuracy of the estimated probabilities $\hat{p}$. Mathematically, the sample size is calculated as:

$$n = \frac{1-p}{pa^2}[\Phi^{-1}(1-\frac{\alpha}{2})]^2,$$

[20] A copy of the survey can be found in the appendix.

where p is the choice proportion of the relevant population, a is the level of allowable deviation as a percentage between $\hat{p}$ and p, the parameter α is the type I error where as an α value of 0.05 indicates that parameter estimates are statistically significant at 95%, and $\Phi^{-1}(1-\frac{\alpha}{2})$ is the inverse cumulative distribution of a standard normal distribution at $\left(1-\frac{\alpha}{2}\right)$. In my experiment, I have 4 tracker options. Assuming they are equally likely to be chosen, the choice proportion $p = 1/4 = 0.25$. The choice of accuracy is somewhat subjective under the rule that the more accurate the estimates are the larger sample is required. If the desired level of accuracy is 30% from the mean ($a = 0.3$), then the required sample size is:

$$n = \frac{(1-0.25)}{0.25 * 0.3^2}\left[\Phi^{-1}\left(1-\frac{0.05}{2}\right)\right]^2 = 65,$$

or approximately the sample I recruited. This is a reasonable sample size in social-networking experiments as samples are necessarily small due to the computational difficulty in estimating with large social weight matrices, and the practical necessity of ensuring that each subject can plausibly assign relational values to all others.

The sample I used for my study was a student sample. Student samples are often used for laboratory experiments (Narayan, Rao and Sanders 2011, Richards, Hamilton and Allender 2014). Although samples from the general population may be more representative of the relevant market in terms of demographic and socioeconomic characteristics, I focus on the behavioral patterns and not the actual WTP for activity trackers, *per se*. That is, even though students may not choose exactly the same trackers as subjects drawn from the general population, any differences in sample composition should not affect how the individuals respond to peers.

In order to provide a sense of what the sample looks like relative to the general population, I provide a set of summary statistics that compares the composition of my sample to the population in Table 4.3. The sample consists of mostly junior and senior business and engineering students. Subjects average 20.59 years of age, relative to the state mean of 37.2. A younger sample is to be expected because it consists entirely of students. Further, 28.6% of my sample is female compared to the state mean of 50.6%, which again is to be expected given that my sample is drawn from colleges that tend to be overrepresented by male students. Regarding ethnicity, my sample contains 46% White, 21% Asian, 16% Hispanic, 2% of Native American, and 7% other races. Compared to the state mean of 57.8% White, 3.4% Asian, 30.3% Hispanic, and 0.3% Native American, White and Hispanic are under-represented while Asian and Native American are over-represented.

Behavioral attributes may also be important determinants of purchase intent for activity trackers. With respect to previous purchase experience, only 17.5% of the sample own an activity tracker. The sample is relatively active with 25.4% working out every day; while the majority (52.4%) work out at least once a week, 11.1% work out once every other week, and the rest work out less frequently. On average, subjects spend $187 annually on sports-related purchases with a standard deviation of $146. The frequency of purchases lies mostly in the category of “once every 3 months”, followed by “once every 6 months” and “once every month”. The average BMI of my sample is 24.39, which is within the normal range of weight/ height ratio. Even though I recorded information on income, students who still live with their parents report household income, and students who have jobs and don’t live with their parents report their own income. The sample shows that the majority (54.3%) has an income of less than $39,000/ year, while the rest has an income of more than $40,000/year. Comparing to the average income of Arizona (2013) at $48,510 and the average income of the nation at $52,250 (Arizona Population Statistics Census Data 2015, https://population.az.gov/), the sample income is below average but representative of a student sample.

Table 4.3 Descriptive Statistics for the Sample

Variable	Definition	Frequency %	Mean	Std. dev
Gender	Gender of participant Female=1; Male=0		0.29	0.455
Age	Age in years		20.59	2.519
Annual household income	Total household income			
	Less than $10,000	33.3		
	$10,000 to $19,999	10.5		
	$20,000 to $29,999	10.5		
	$30,000 to $39,999	0		
	$40,000 to $49,999	3.5		
	$50,000 to $59,999	8.8		
	$60,000 to $69,999	12.3		
	$70,000 to $79,999	5.3		
	$80,000 to $89,999	0		
	$90,000 to $99,999	3.5		
	$100,000 to $149,999	7.0		
	More than $150,000	5.3		
Workout frequency	How often do you work out?			
	Every day=5	25.4		
	At least once a week=4	52.4		
	Once every other week=3	11.1		
	Once a month=2	6.3		

	Once a few months or less often=1	4.8		
Purchase_freq (purchase frequency)	How often do you purchase sports goods?			
	At least once a week=5	4.8		
	Once a month=4	17.5		
	Once every three months=3	34.9		
	Once every six month=2	27.0		
	Once a year or less often=1	15.9		
Purchase_expPurchase expenditure	How much money do you spend sports goods?		187.30	146.179
Tracker ownership	Yes=1; No=0		0.17	0.383
BMI			24.396	6.79
Ethnicity	White		0.46	0.502
	Hispanic		0.16	0.373
	Native American		0.02	0.128
	Asian		0.21	0.413
	Other		0.07	0.250

In the next section I estimate attribute-preferences by calibrating two models: a random parameter model to derive subjects' willingness to pay, and a spatial model to identify peer effects. Attribute preferences can only be inferred in a CBC experiment by econometrically estimating their marginal value in a formal, utility-theoretic framework. Peer effects are derived then by studying the driving factors of willingness to pay.

4.5 Econometric Model of Preference Revision

WTP, in the sense of consumer demand, is the value a consumer places on a higher utility level relative to a lower utility level given a budget constraint[21]. Estimating

[21] In other words, a Hicksian surplus.

the WTP for consumers is a common practice in evaluating novel products or changes of quality in existing goods (Lusk and Hudson 2004, Lusk, Roosen and Fox, 2003, Nalley *et al.*, 2004). For example, Tonsor, Olynk, and Wolf (2009) assess consumer willingness to pay (WTP) for animal welfare attributes in meat products. Narayan, Rao and Sanders (2011) use WTP to reveal the preference revision of e-book readers for additional functions. For the purpose of studying activity tracker choices and deriving individual preferences, I calculate the individual WTP and utilize a set of spatial models to analyze preference revision and peer effects. In the following section, I present the discrete models that derive WTP first, followed by the spatial models that analyze the preference revisions.

4.5.1 Choice modeling

The objective of the experiment is to elicit changes in preferences due to social interaction. Because I am interested in preference changes from stage one to stage two as it relates to social interaction, my econometric model estimates preferences in stage one, and then testis for the significant preference revision between stages one and two. I do this in two ways. First, I test for changes in attribute valuation from stage one to stage two due to the treatment effect of interacting with others in a direct way. Second, I calculate WTP for activity trackers, and test whether changes in WTP depend upon the social influence from stage one to stage two, moderated by the degree of social relationship between each subject.

Consistent with the data generated by my CBC experiment, the core of the econometric model consists of a discrete-choice model, which is particularly adept at estimating marginal attribute valuations (Train 2003, Hensher, Rose and Greene 2005).

This point is important because discrete choice models, particularly those of the logit form used here, provide closed-form choice probability expressions that are useful in calculating choice probabilities under a range of peer-influence assumptions.

When using discrete choice models, an individual's utility is considered a random variable either because the researcher has incomplete information, or there is unobserved heterogeneity in individual preferences (Manski,1977). Formally, let the i th consumer's utility from choosing alternative j be given by:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \qquad (4.1),$$

where V_{ij} is the deterministic component of the utility function determined by the interested activity tracker attributes and ε_{ij} is the random component. Assuming V_{ij} is linear in parameters, the form of the utility function for alternative j can be expressed as:

$$V_j = \sum_i \beta_i x_{ij} = \beta_1 x_{j1} + \beta_2 x_{j2} + \cdots + \beta_n x_{jn} \qquad (4.2)$$

where x_{jn} is the full vector of explanatory variables that are observed by the analyst, including attributes of the alternatives, , and variables that describe treatment and stage effects, and β_i is a vector of parameter estimates associated with x_{jn} The estimated β values in this exercise are of particular importance because they measure the marginal utility of each tracker attribute. The explanatory x_{jn} variables are listed in Table 4.4. The variable "None" represents the "none of these" option in the consumer's choice set, and serves as the "outside option" in discrete-choice modeling terminology (None = 1 if "none of these" option is selected, None = 0 otherwise). Garmin, Jawbone, Nike, and Fitbit are dummy variables that represent the four different brands, while Clip-on, Wristband, and Watch are dummy variables that represent the different activity tracker designs. Different functions are represented as dummy variables that capture the ability to record calories only (Cal), sleep patterns (Sleep), text and email messages (Msg), and recording workout routes with the aid of a Global Positioning Satellite (GPS). The variable Price captures the price of the respective alternative. Besides the attributes, I also included interaction variables to capture the difference between stages as well as treatment effects. Stage interactions variables are the product of a stage binary variable and all attributes, indicating whether there exist significant differences between stages. Three-way interactions of stage and treatment represent the attributes effect at stage 2 among treatment groups. These interactions will help reveal whether there exist treatment effect in the second stage.

Table 4.4 List of Variables

Variable Name	Meaning
Garmin	The brand is Garmin
Nike	The brand is Nike
Fitbit	The brand is Fitbit
Jawbone	The brand is Jawbone
Clipon	Design is clip on
Watch	Design is watch
Wristband	Design is wristband
Cal	Function is recording calories only
Sleep	Has additional function of recording sleeping patterns
GPS	Has additional function of recording workout route
MSG	Has additional function of receiving messages
Price	Price of the tracker
S_garmin	The brand is Garmin in the second stage: Garmin*stage 2
S_nike	The brand is Nike in the second stage: Jawbone*stage 2
S_jawbone	The brand is Jawbone in the second stage: Jawbone*stage 2
S_watch	The design is watch in the second stage: watch*stage 2
S_wristband	The design is wristband in the second stage: wristband*stage 2
S_sleep	Has additional sleeping function in the second stage: Sleep*stage 2
S_gps	Has additional GPS function in the second stage: GPS*stage2
S_msg	Has message function in the second stage: MSG*stage 2
ST_garmin	Three-way interaction: Garmin*Treat*Stage2

ST_nike	Three-way interaction: Nike*Treat*Stage2
ST_jawbone	Three-way interaction: Jawbone*Treat*Stage2
ST_watch	Three-way interaction: watch*Treat*Stage2
ST_wristband	Three-way interaction: Wristband*Treat*Stage2
ST_sleep	Three-way interaction: Sleep*Treat*Stage2
ST_gps	Three-way interaction: GPS*Treat*Stage2
ST_msg	Three-way interaction: msg*Treat*Stage2
None[b]	Choose the "none of these function"

Assuming the random error term in equation (4.1) is distributed Type I Extreme Value (EV), the probability of choosing option j over option *k* is:

$$Prob\{j\ is\ chosen\} = Prob\{V_{ij} + \varepsilon_{ij} \geq V_{ik} + \varepsilon_{ik}\} = \frac{e^{V_{ij}}}{\sum_{k \in C} e^{V_{ik}}} \quad (4.3),$$

where choice *j* is chosen over choice *k* if the overall utility for choice *j* is greater than the utility of choice *k*.

One well-understood problem with the logit framework is that it implies that the ε_{ij} are independent and identically distributed (IID) across individuals and alternatives. The IID assumption is restrictive in that it does not allow for the error components of different alternatives to be correlated (Hensher, Rose and Greene 2005). Therefore, I relax the iid assumption by using a random-parameter logit (RPL), or mixed logit, model.

Importantly, a RPL model allows taste parameters to vary randomly in the population. Formally, each marginal attribute value is written as (Hensher, Rose and Greene 2005):

$$\beta_{ik} = \beta + \delta'_k z_q + \eta_{ik} \quad (4.4),$$

where η_{ik} is a random term that is distributed randomly over individuals. The random term can assume a range of distributions, depending on the choice environment. In this model, z_q is observed data specific to the individual with q random variables and, η_{ik} denotes a vector of *k* random components in the set of utility functions in addition to the J random elements in ε_{ij}. The error term can assume different distributional forms such as normal, lognormal and triangular. For my study I choose triangular distribution for Price and normal for all other random parameters[22]. Triangular distribution guarantees a positive price estimate therefore a positive WTP, while normal distribution is the most common practice for random variables. For a give value of η_q, the conditional probability L_{jq} of choosing option j is the following (given the remaining error term is IID):

$$L_{jq}(\beta_q|X_q, \eta_q) = \exp(\beta'_q x_{jq}) / \sum_j \exp(\beta'_q x_{jq}) \tag{4.5}$$

Equation (4.5) is the simple MNL model, but for each sampled individual, there is additional information defined by η_q. The unconditional choice probability is the expected value of the logit probability over all the possible values of β_q, that is, integrated over the values of β, weighted by the density of β_q. The probability is presented as:

$$P_{iq}(X_i, z_i) = \int L_{jq}(\beta_q|X_q, \eta_q) f(\eta_q|z_q) d\,\eta_q, \qquad (4.6),$$

where $\beta_q = \beta + \Delta Z_q + \eta_q$. Thus, the unconditional probability that individual q will choose alternative j given the specific characteristics of their choice set and the underlying model parameters is equal to the expected value of the conditional probability as it ranges over the possible values of β_q. The random variation in β_q is induced by the random vector η_q.

[22] Hensher, Rose and Greene (2005) suggest specifying the parameters associated with each attribute (including price) as random (see Hensher, Rose and Greene 2005, pg.618) because inter-alternative error correlation could be confounded with unobserved preferences if the latter is not explicitly taken into account (Daniels and Hensher 2000, Bhat and Castelar2003).

The choice probabilities associated with the RPL will not exhibit the same IIA property as the fixed-coefficient logit, and may yield different substitution patterns by appropriate specification of $f(\eta_q|z_q)$. Flexible substitution is introduced through the random parameters, specifying each element of β_q associated with an attribute of an alternative as having a mean, a standard deviation, and possibly a measure of correlation with another random parameter. By allowing marginal attribute valuations to vary across sample subjects, I am able to determine how preferences are influenced by exposure to the choices of others.

Among the variables included in the indirect utility function, I allowed the marginal utility of income (price parameter) to vary randomly with a triangular distribution. A triangular distribution is highly desirable because it binds the parameter on (-1,1). Allowing the marginal utility of income to vary randomly is a common practice, and reasonable as this parameter governs price-response and price-response is driven by behavioral attributes of the household, many of which are unobserved (Hensher, Rose and Greene 2005, Banerjee, Martin and Hudson 2006, Lusk and Norwood 2009, Tonsor, Olynk, and Wolf 2009).

Recall that the objective of this study is to reveal individual preferences, and how they are revised through peer influences. I measure preference revision through the WTP, which is the amount of money a subject is willing to pay for a unit change for a particular attribute. In the current model, I define the mean price parameter as β_1, and an attribute whose parameter is normally distributed with mean β_2 and standard deviation σ_2. Willingness to pay is calculated as:

$$\text{WTP} = -\frac{\beta_2}{\beta_1} \quad (4.7),$$

where the WTP for the attribute is distributed normally with mean $\frac{\beta_2}{\beta_1}$ and standard deviation $\frac{\sigma_2}{\beta_1}$ (Hensher, Rose and Greene 2005). Because WTP measures are calculated as the ratios of two parameters, they are sensitive to the range of each attribute level used in the estimation of both parameters. I then use the individual- and attribute-specific WTP calculated above and take difference between the second stage and the first stage WTP values to calculate a measure of preference revision:

$$\Delta WTP_J = WTP_{J,Stage2} - WTP_{J,Stage1}.$$

Differences between first- and second-stage valuations may be positive or negative, depending on the nature of the information received between the two sessions. For current purposes, however, I am more interested in how preferences are moderated by social interaction than the direction of change. For this purpose, I estimate using a spatial econometric approach that I describe in the next section.

4.5.2. Spatial Models

Spatial models are used to estimate preferences in a social environment because they are non-linear in structure and account for simultaneous interactions among individuals through the social weight matrix (Anselin 2002, Yang and Allenby 2005, Richards, Hamilton and Allender 2014). Spatial models differ from traditional linear-in-mean models in that they address the need for a multidimensional relationship between consumers through the weight matrix (Lee 2004). Moreover, the natural exclusion restrictions implied by the social network structure ensure the separate identification of endogenous and contextual peer effects (Lin 2014). In particular, for the linear-in-means model, peers' outcomes are measured by group mean outcomes, and peers' characteristics are captured by group mean characteristics. Both measurements are group-specific and constant for all members in the same group. The consequence is that these two terms are linearly dependent, and the endogenous effects cannot be separated from the contextual effects.

There are a multitude of different forms of spatial model, but I focus on two types: a Spatial Autoregressive Model (SAR, or "spatial lag model" as it is referred to by Anselin (2002)) and a Spatial Error Model (SEM). The classic SAR model departs from the Manski (1993) model by measuring peer variables as the weighted averages of observed peer outcomes and characteristics instead of group expectations. Both peer outcomes and peer characteristics are specific to the individual and vary across group members. The SEM model instead captures spatial patterns in the error term; therefore, it accounts for the unobserved heterogeneity in consumer tastes. SEMs treat spatial correlation primarily as a nuisance, similar to how statistical approaches treat temporal serial correlation. This approach generally focuses on estimating the parameters for the independent variables of interest in the systematic part of the model, and essentially disregards the possibility that the observed correlation may reflect something meaningful about the data generation process (Ward and Gleditsch 2007). When peer outcomes are caused by an unobserved correlated effect, the SEM will capture it by regressing a spatial weight matrix on the residuals.

A SAR-SEM model is a spatial model that incorporates both SAR and SEM features. The SAR-SEM model used here similar to the one used by Lin (2014) in that I use a spatial weight matrix that represents the actual relationship among members, but differs in that I relax the group-specific effect and instead use a binary variable to indicate differences between treatment and control groups. This is critical to my approach, because peer recommendations are only introduced in the treatment groups. If there indeed is a pure peer effect then the group fixed specific effect will capture it. With this assumption, the model is written as:

$$y_{ik} = \alpha + \rho \sum_{j=1}^{n} w1_{ij} y_{jh} + x_{ik}\beta + \sum_{j=1}^{n} w_{ij}\, h_{ik}\gamma + G\delta + u_{ik} \qquad (4.9),$$

and:

$$u_{ik} = \lambda \sum w2_{ik}\, u_{ik} + \varepsilon_{ik}, \qquad (4.10),$$

where y_{ik} is individual WTP revision for attribute k, ΔWTP_J. Variables x_{ik} are individual characteristics related to the purchase of activity trackers; h_{ik} are individual characteristics about *i*'s background averaged over the group; w_{ij} is the *ij* element of a row-standardized, zero diagonal weight matrix that captures the network structure where *i* and j are different subjects, and *G* is a fixed-group effect with a binary indicator for treatment-group membership. The error terms (eq. 4.10) in the model, u_{ik}, follow an SEM process, which captures the unobserved effects that vary within the group and thus cannot be captured by the group fixed effects; and $\varepsilon_{ik}$is an idiosyncratic error term. This specification is the most general form of spatial model, termed the SAR-SEM model by LeSage (1998), because it captures both direct spatial effects through the SAR term and indirect effects, through unobservable elements, in the SEM term.

In the SAR-SEM social model, outcomes for individuals from the same group are correlated in multiple ways. First, the parameter ρ measures the endogenous effect of others' behavior on each agent's WTP, second, γ captures the contextual effect, and, third, the group fixed effect is represented by δ, capturing the common factors that affect all group members. Finally, the possibility that subjects' choices are correlated through spatially dependent unobservable is captured by λ in the error term. Intuitively, the parameter ρ estimates the presence of a spatial lag effect, or that consumer's preferences are influenced by her peers. The value of ρ is bounded by 0 and 1. A parameter close to 1 indicates greater influence, and a parameter of 0 means there is no influence at all. A negative value of ρ indicates a consumer is negatively influenced by her peers, whereas a positive value of ρ indicates the consumer follows her peers' decisions. Estimates of γ indicate contextual effects, which are the factors related to the common environment such as education, race, income, age, and gender. Group effects are estimated with the δ parameter, which is interpreted as the influence of peer recommendations introduced only in the treatment groups. In this regard it measures the difference in preferences between the control and treatment groups. Finally, after accounting for the peer, contextual, and group fixed effects, λ captures any "left-over" unobserved effects that exist in the data. In the rest of this section, I discuss how the peer effects are identified in the SAR-SEM model in econometric terms – an identification strategy that relies critically on the nature of the spatial weight matrix.

At the core of any spatial model is a social weight matrix. The structure of the social weight matrix, $\boldsymbol{W}$, is essential to estimating peer effects with this model. A social weight matrix is a $n \times n$ positive matrix where n is the number of members, $\boldsymbol{W}$, through which the ''neighborhood set'' is specified for each observation. An observation appears both as a row and column, with non-zero matrix elements w_{ij} indicating the strength of the peer relationship between participants (row) i and (column) j. By convention, self-neighbors are excluded, such that the diagonal elements $w_{ii} = 0$. Also, the weight matrix is typically row-standardized, with weights $w_{ij}^{s} = w_{ij}/\sum_j w_{ij}$. Row-standardization means that pre-multiplying another vector creates an average of the neighboring values in the spatial lag operator (Anselin, 1988b). For this study, I use two set of spatial weight matrix: $\boldsymbol{W}_{CLOSENESS}$ that describes the tie strength, and $\boldsymbol{W}_{RELIABILITY}$ that describes the source expertise of network members.

These two weight matrices essentially represent two different mechanisms through which preferences may be revised through social interaction. The ***W*** (closeness) matrix captures tie strength in which preferences are revised through established social distance between individuals. On the other hand, ***W*** (reliability) captures "source credibility" in that revisions are moderated by the extent of credence individual *i* lends to individual's *j*'s comments regarding the product.

For estimation purposes I define ***W*** in terms of a general weight matrix and rewrite Eq. (4.9) in matrix notation:

$$\boldsymbol{Y} = \rho \boldsymbol{W} y + \boldsymbol{X}\beta + \boldsymbol{W}\boldsymbol{H}\gamma + \boldsymbol{G}\delta + u \qquad (4.11),$$

where ***Y*** is a vector of individual differences in WTP regarding a specific attribute, ***X*** is a vector of individual characteristics that will influence the purchase of activity trackers, which includes purchase frequency, workout frequency, purchase amount in dollars, and whether the subject owns a tracker (Own). I expect that purchase frequency, workout frequency and purchase amount in dollars are positively related to WTP because these variables measure the extent of physical activity and expenditure on sporting goods that should be positively related to the WTP for an activity tracker.

The vector $\boldsymbol{H}$ measures characteristics of the reference group, including age, income, gender, and whether the subject is white (White) in this vector. To find the contextual effect, I pre-multiply the vector $\boldsymbol{H}$ that contains information about each subject's background and environment with the row standardized weight matrix $\boldsymbol{W}$. The vector $\boldsymbol{G}$ represents membership in the treatment group and is noted as "1" for treatment groups and "0" for control groups. Table 4.5 shows a list of variables used in this model.

Table 4.5 Variables included in the SAR

Variables	Meaning
Individual characteristics	
Purchase_freq	Purchase frequency of sporting good
Purchase	$ spent on purchasing sporting goods annually
Workout	Workout frequency
Tracker	Whether the subject owns a tracker
Contextual effects	
Age	Age of the subject
Gender	Gender of the subject
Income	Household income
White	If the subject is white

Spatial models are rarely estimated in the form given in (4.11), however, due to the obvious endogeneity of the lagged peer-effect variable. Instead, reduced-form expressions are derived, and estimated. Specifically, the dependent variable $\boldsymbol{Y}$ appears on both sides of (4.11). Clearly, this variable is not exogenous, so I re-write the structural form of 4.11 in a reduced form in order to solve for Y:

$$\begin{aligned}\boldsymbol{Y} = (1-\rho\boldsymbol{W})^{-1}\boldsymbol{X}\beta + (1-\rho\boldsymbol{W})^{-1}\boldsymbol{WZ}\gamma + (1-\rho\boldsymbol{W})^{-1}\boldsymbol{G}\delta \\ + (1-\rho\boldsymbol{W})^{-1}\lambda\boldsymbol{W}u \quad + (1-\rho\boldsymbol{W})^{-1}\epsilon\end{aligned} \qquad (4.12),$$

where $(1-\rho\boldsymbol{W})^{-1}$ is defined as the "inverse Leontief matrix." Writing the model in reduced form highlights the value of using a spatial approach to estimating models of social influence as the inverse Leontief matrix is often described as a "spatial filter". Spatial filtering essentially means that the econometric procedure extracts that part of the variation in the endogenous variable that is due solely to relationships with other spatial observations. What is left, therefore, has the spatial effects removed, or "filtered" out. More formally, this matrix is a full inverse, which yields an infinite series that involves all neighbors: $1 + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + \cdots \rho^n\boldsymbol{W}^n$. This means that each neighbor is correlated with every other neighbor, but the correlation decays with the order of contiguity (the powers of $\boldsymbol{W}$ in the series expansion). Higher powers of the weight matrix ($\boldsymbol{W}^n$) reflect neighbor sets in more remote contiguity (nth neighbor). This illustrates the global nature of the spatial multiplier effect in the spatial lag model (Anselin 2002). Specifically, if a unit change were introduced in a given explanatory variable X_k, the effect on y would amount to $[1/(1-\rho)\beta_k]$. More generally, for any vector of changes in a given explanatory variable, $\Delta \mathrm{x}_k$, the resulting spatial pattern of changes in the dependent variable is a function of the spatial filter and the change of given explanatory variables:

$$\Delta y = (1 - \rho\boldsymbol{W})^{-1}\Delta \mathrm{x}_k \beta_k \qquad (4.13).$$

This expression conveys the intuition that changes in preferences in a social environment derive from two sources: a spatial component and an explanatory variable component.

Indeed, the global nature of social interactions is apparent through the reduced form. For instance, the spatial lag term $\rho\boldsymbol{W}$ for observation j is correlated with its own error u_j, and with all other errors in the system, which accounts for spatial correlation among the explanatory variables and peer effects. Thus, the estimate of β (obtained after spatially filtering the dependent variable y) is a consistent estimate of the marginal value of product attributes ($\boldsymbol{X}$ on $\boldsymbol{Y}$). This means that after spatially filtering out the network effect by multiplying the βs with $(1 - \rho\boldsymbol{W})^{-1}$, the estimated βs are truly the individual impact of product attributes without confounding social effects, or perpetual affinity because it shows how a subject's personal preference towards activity trackers influence her WTP. Further, the estimate of ρ is a consistent estimate of the peer effect because the SAR process accounts for the global nature of peer influence, that is, to which degree a subject is influenced by all her peers at once.

Estimation of the SAR-SEM model is difficult, yet the consequences of ignoring spatial dependence in models can be substantial. If a causal relationship of the dependent variables among peers does exist, but the model is estimated without the spatial autoregressive term, then a significant explanatory variable has been omitted, and the estimated coefficients will be biased and inconsistent (Kalnins 2003). On the other hand, if there is unobserved correlation among the error terms due to spatial dependence, then a SEM is required to obtain unbiased and consistent estimates. In the next section, I address how the model is estimated with a two-stage instrumental variable (IV) method. Moreover I present the specification tests that evaluate spatial lag dependence, spatial error dependence and both.

4.5.3 Estimation

The estimation problems associated with spatial regression models are distinct for the spatial lag and spatial error case (Anselin 2006). Spatial error models are special cases in which the error is non-spherical, or violate the fundamental assumptions of OLS estimation. In other words, if there exists unobserved correlation effect that causes consumers to make similar decision then a spatial error test should detect significant spatial dependence in the error term. On the other hand, the inclusion of a spatially lagged dependent variable results in a form of endogeneity. A classic solution to the endogeneity problem is to use instrumental variables. Kelejian and Robinson (1993) suggest the use of a subset of columns from {X, WX, W^2X^2,W^3X^3, . . . } as instruments. Specifically, the optimal instruments are:

$$\boldsymbol{Q}=I-\lambda\, \boldsymbol{W_2}\, [\boldsymbol{X}, \boldsymbol{W_1}\, (I-\rho\, \boldsymbol{W_1})- \boldsymbol{W_1} x\beta\,] \qquad (4.14),$$

where $\boldsymbol{Q}$ *(N × q)* is a vector of instrumental variables. In the case of peer outcomes, I use the lag of all explanatory variables as my instruments. This is a common practice for such models (Anselin 2009) and essentially adds to the explanatory power of peers. For the simplicity of notation, I write Equation 4.11 as:

$$\boldsymbol{Y} = \boldsymbol{Z}\xi + u \qquad (4.15),$$

where **Z** = [$\boldsymbol{W}_1$y, **X**, **WH**, **G**] and ξ = [ρ, β, γ, δ]. The generalized spatial two-stage least squares estimator developed in Kelejian and Prucha (1998) consists of three steps. The first step is a spatial two-stage least squares estimation. The predicted value of $\boldsymbol{Z}$ in a regression on the instruments is obtained as:

$$\widehat{Z} = \boldsymbol{Q}(\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'\, \boldsymbol{W}\boldsymbol{Z}, \qquad (4.16),$$

The instrument $\widehat{Z}$ replaces $\boldsymbol{Z}$ in the second stage, resulting in the spatial two-stage least squares estimator:

$$\widehat{\xi_{S2SLS}} = [\widehat{Z}'\widehat{Z}]^{\boldsymbol{-1}}\, \widehat{Z}'\xi, \qquad (4.17),$$

or in full,

$$\widehat{\xi_{S2SLS}} = [\boldsymbol{Z}'\boldsymbol{Q}(\boldsymbol{Q}'\boldsymbol{Q})^{\boldsymbol{-1}}\boldsymbol{Q}'\boldsymbol{Z}]^{-1}\, \boldsymbol{Z}'\boldsymbol{Q}(\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'\boldsymbol{Z}\xi \qquad (4.18).$$

With asymptotic covariance matrix given by:

$$AsyVar\left[\widehat{\xi_{S2SLS}}\right] = \hat{\sigma}^2[\boldsymbol{Z}'\boldsymbol{Q}(\boldsymbol{Q}'\boldsymbol{Q})^{\boldsymbol{-1}}\boldsymbol{Q}'\boldsymbol{Z}]^{-1} \qquad (4.19).$$

The solution of the system by nonlinear least squares yields a consistent estimate $\hat{\lambda}$ for the autoregressive error parameter (Anselin 2006).

The purpose of spatial econometrics is to determine whether any spatial relationship of the variables is merely random or responds to a pattern of spatial dependence. Specification testing is necessary to find the spatial patterns in any given data set. Each specification test is constructed with a specific alternative in mind, so that the test consists of a test of restrictions on the parameters of a model that includes spatial dependence, such as a spatial error model or a spatial lag model. The literature on specification tests against spatial correlation in cross-sectional regression is by now quite extensive (Anselin and Bera, 1998; Anselin, 2001a; Florax and de Graaff, 2004). The most commonly used approach under maximum likelihood estimation is Lagrange Multiplier (or Rao Score) tests, which are based on the slope of likelihood function, or the "score" function. In particular, tests against the presence of spatial correlation are very important, as ignoring spatial correlation when it is present may lead to biased and inconsistent estimates of the model parameters, or inefficient estimates and biased t-test statistics. In the following section, I present the commonly used specification tests to detect spatial lag dependence and error dependence, organized as tests against spatial autocorrelation, tests based on the Maximum Likelihood principle, and tests against multiple sources of misspecification.

The $n \times n$ spatial weight matrices $\boldsymbol{W}$ consist of exogenously specified elements (that capture the neighbor relations of observations *i* and *j*) in order to identify peer effects. A Lagrange Multiplier tests the residuals of an ML estimate of the null model that includes a single $\boldsymbol{W}$ matrix. For my purposes, the specification is to estimate ρs for $\boldsymbol{W}$s, which is the matrix of peer relationships. The residuals can then be used for a test of whether the coefficient ***X***s of ***W***s is significant. The null is defined as the classic linear regression model. Mathematically, the LM test statistic is Chi-square distributed and is written:

$$LM_\rho = \{\frac{e'Wy}{\frac{e'e}{n}}\}^2/D, \tag{4.20),}$$

where the first term is the residual sum of squares on $\boldsymbol{X}$, and the denominator is a scaling factor that is based on the weight matrix and estimates of the OLS. While e is the OLS residuals and the denominator D is:

$$D = \left[\frac{(\boldsymbol{WX}\hat{\beta})'[I - \boldsymbol{X}(\boldsymbol{X'X})^{-1}\boldsymbol{X'}](\boldsymbol{WX}\hat{\beta})}{\hat{\sigma}^2}\right] + tr(\boldsymbol{W'W} + \boldsymbol{WW}) \tag{4.21),}$$

where the estimates for $\hat{\beta}$ and $\hat{\sigma}^2$ are from OLS. The test statistic is asymptotically distributed as $\chi^2(1)$. For this test, the null hypothesis is H_0: $\rho = 0$, so the alternative is an OLS model. Failing to reject the null hypothesis indicates consumers do not depend on their peers for opinions, while rejecting the null hypothesis indicates the existence of peer effects. Basically, the LM is testing the slope of the log-likelihood function when there exists spatial lag against the log-likelihood when there is no spatial lag. If the slope is significant, then I reject the null hypothesis of OLS in favor of a spatial lag specification.

The point of departure for a LM test for spatial error autocorrelation is that it tests the unobserved correlation in residuals that might cause consumers to show similar preferences. Therefore, instead of regressing consumer outcomes (***Y***) on peer relations, the LM error test investigates whether the unexplained residual displays any sort of spatial correlation. Again, the null hypothesis is an OLS with H_0: $\lambda = 0$. Mathematically, the test is written as:

$$LM_\lambda = \frac{[\frac{e'\boldsymbol{W}e}{e'e/n}]^2}{tr[\boldsymbol{W}'\boldsymbol{W} + \boldsymbol{W}\boldsymbol{W}]}, \qquad (4.22),$$

where e is a $n \times 1$ vector of OLS residuals and is asymptotically distributed as $\chi^2(1)$. Failing to reject the null hypothesis indicates peer relationships and the explanatory variables explain all social effects, while rejecting the null hypothesis means there are unexplained effects that correlate with peer relationships.

Besides testing for peer effects and unobserved effects separately, Anselin (1998b) provides a joint test where the null hypothesis is H_0: $\rho = \lambda = 0$. This LM test is not simply a summation of the two statistics above, but takes on a more complicated form given by:

$$LM_{\rho\lambda} = \frac{d_\lambda^2 D + d_\rho^2 T_{22} - 2 d_\lambda d_\rho T_{12}}{D T_{22} - T_{12}^2} \qquad (4.23),$$

where $d_\lambda = \frac{e'We}{e'e/n}$, $d_\rho = \frac{e'Wy}{\frac{e'e}{n}}$ and $T_{ij} = tr[W_i W_j + W_j' W_j]$. Intuitively, if both spatial lag dependence and spatial error dependence are significant, then the joint test statistic should be significant too. I apply each of these tests before choosing the preferred specification in interpreting the experimental data below.

4.6 Results

In the first part of this section, I present the results of the discrete choice model, with which I derive attribute specific WTP and answer the question of which attributes significantly influence activity tracker choices. In the part of this section, I identify specific social effects, namely peer effect, contextual effect, group effect and unobserved correlation through a set of spatial models. In addition, I address hypotheses regarding peer recommendations, and provide a set of implications for marketing practice, and the economics of social learning.

4.6.1 Multinomial Logit Results

To start with, I provide estimates from a fixed-coefficient MNL model in order to serve as a starting point for interpreting the random parameter logit model. Table 4.6 reports the MNL estimates obtained with the experimental data. Based on the estimates shown in table 4.6, I find that all the variables except for Jawbone are significant. This means that the variables included in the model explain the probability of activity tracker choices well. Among brands, these estimates show that all brands are significant compared to the baseline (Fitbit). Therefore, this estimate is interpreted as showing that subjects are more likely to choose Garmin and Nike compared to choosing the Fitibt. Considering that the Jawbone coefficient was not found to be significant, this means that the brand is not influencing consumers' decision making.

Table 4.6 Estimated parameters for MNL and RPL.

	MNL		RPL			
Variable	Mean	t-ratio	Mean	t-ratio	Standard Deviation	t-ratio
Garmin	0.3739*	2.42	0.3886 *	2.43	0.0155	0.14
Nike	0.6554*	4.45	0.6154*	3.84	0.4808*	4.36
Jawbone	-0.0272	-0.18	-0.1579	-0.96	0.4720*	4.32
Watch	1.3910*	10.05	1.3855*	9.11	0.5476*	4.50
Wristband	1.2080*	9.05	1.2767*	9.24	0.1421	1.28
Sleep	1.3831*	9.23	1.4845*	9.51	0.0461	0.41
GPS	1.4860*	9.24	1.6474*	9.55	0.4801*	4.53
MSG	0.9783*	6.32	1.0075*	6.24	0.2695*	2.50
Price	-0.0115*	-11.71	-0.0170*	-13.36	0.0170*	13.36

S_garmin	0.9020*	3.57	0.9377*	3.65
S_nike	0.3739	1.29	0.3590	1.19
S_jawbone	0.5433*	2.14	0.6227*	2.34
S_watch	-0.7648	-3.44	-0.7867*	-3.34
S_wristband	-1.4356	-6.04	-1.5052*	-6.21
S_sleep	-1.2693*	-5.15	-1.3634*	-5.33
S_gps	-1.5765*	-5.98	-1.7418*	-6.29
S_msg	-0.7882*	-3.16	-0.8646*	-3.35
ST_garmin	0.1723	0.62	0.1498	0.53
ST_nike	0.29596	0.89	0.2831	0.81
ST_jawbone	0.4249	1.50	0.3257	1.07
ST_watch	0.2153	0.92	0.1821	0.71
ST_wristband	0.1331	0.52	0.0853	0.32
ST_sleep	0.2552	0.93	0.2469	0.87
ST_gps	0.2639	0.94	0.3236	1.06
ST_msg	-0.0373	-0.14	-0.0339	-0.12
None of these	0.8127*	-11.71	0.5179*	2.72
Summary statistics				
Number of observations		1512		1512
Number of participants		63		63
LL^c mixed logit		-2095.8949		-1990.4450
LR				105.4499

*significant at 95% level.

In order to get attribute and individual specific WTP, a Random Parameter Model (RPL) is needed to account for the randomness in tastes. To do so I let all variables used in the MNL vary randomly according to a known distribution and perform a LR test between the RPL and the MNL. Rejecting the MNL in favor of the RPL will indicate that consumers preferences are indeed heterogeneous and vary based on the attributes. Specification tests for discrete choice models typically involve comparing the estimated model with a null alternative using likelihood ratio (LR) tests. [23] Estimating two models and comparing the fit of one model (log-likelihood value, LL) to the fit of the other (LL) perform the LR test. Log-likelihood function provided values to assess model fit where the bigger the log-likelihood value (LL) is the better. A LR test compares the performance of the MNL and RPL models by taking the difference between the goodness of fit of two models and compare it to a chi-square distribution:

$$LR = -2(LL_{MNL} - LL_{RPL}) \sim \chi^2{}_{number\ of\ variables},$$

where the LR value greater than a critical value indicates that the RPL performs better than the MNL. The log-likelihood (LL) of the MNL model is -2095.8949 whereas the LL for the RPL model is -1990.4450. Using the Log-likelihood ratio to test the performance of the unrestricted MNL model I find a LR test statistic of 105.4499, which is greater than the critical Chi-square statistic with nine degree of freedom of 16.919[24]. Therefore, in comparison to the MNL model, I am able to conclude on the evidence provided that the fitted RPL model is preferred.

4.6.2 RPL Results

As discussed above, the RPL model is able to capture unobserved heterogeneity, so should provide a better representation of not only the mean parameter estimates, but peer effects as well. Heterogeneity of consumer taste is introduced by allowing variables to be randomly distributed according to a triangular distribution. Technically, any and all

---

[23] The log-likelihood ratio test, which follows a χ2 distribution with *k* degrees of freedom (Wooldridge, 2002) is $-2(LL_i - LL_j)$, where LL_j is the unrestricted pooled sample log-likelihood value and LL_i denote the log-likelihood values for the constant only model. *k* is the number of additional variables compared to the constant only model.

[24] 16.919 is the critical value at 0.95 with 9 degree of freedom, the number of random parameters. The degrees of freedom come from the estimation of variances of random parameters.

of the parameters estimated in the utility model can be regarded as random parameter estimates. Following Hensher, Rose and Greene (2005, pg. 618), I first allow all variables to be random according to a triangular distribution, and then remove the variables that are not statistically significant according to the Wald statistics and the LR test. I define the marginal utility of income with respect to price in terms of a triangular distribution because it binds the parameter dispersion on (-1, 1). Because of this, a triangular distribution on price guarantees positive price thus positive WTP estimates. After selecting the variables for which unobserved heterogeneity appeared to be the most important, I chose Price, Nike, Watch, Wristband, Sleep, GPS and Msg. Among these variables, the scale parameters (standard deviation) are statistically significant for Price, Nike, Watch, GPS and MSG. These results suggest that the random parameter specification is indeed appropriate in these data.

I estimated the extent of preference revision, and test for the effect of peer influence, by comparing marginal attribute valuations between the first and second stages of the experiment. I pooled the data from stage 1 and stage 2 together, and estimate the stage 2 effects by multiplying each variable by a stage 2 indicator variable. If the stage 2 variables are statistically significant, then I can conclude that subjects revise their preferences based on interactions with others in their group. If there is significant revision, then, the parameter estimate for stage 2 is obtained by adding the parameter estimate from stage 1 with the interaction variable. For example, the estimate for Garmin in the second stage is:

$$\beta_{\text{Garmin}} + \beta_{\text{S\_Garmin}} = 0.3739 + 0.9020 = 1.2817$$

These estimates are shown in Table 4.6. Consumers had a higher parameter estimate for all brands after exposure to input from others. However, the parameter values for design (Wristband, Watch) and functions (Sleep, GPS, MSG) were reduced from the first to the second stage. The fact that all the estimated marginal values for the tracker attributes were statistically significant means that subjects clearly revised their preferences between the first and second stages. This notion provides supports to my first hypothesis. More specifically, a positive estimate for the Garmin variable (S_garmin) means that, on average, respondents tended to prefer the Garmin brand, while a negative estimate of the GPS variable (S_gps) means that, during the second stage, respondents were less likely to choose a tracker with the GPS function.

Besides the two-way interactions with stage, I also include a set of three-way interactions with both stage and treatment. These interactions indicate the marginal effects of attributes specific to the treatment group in the second stage. The RPL results show no significant estimates among these interactions, meaning that the RPL does not show significant treatment effect on preferences. The fact that Jawbone was not significant in the first stage, but was significant in the second stage, means that subjects revised their preferences to choose the Jawbone brand between the first and second stages. Whether this preference revision was due to information from specific members of the group, however, requires an econometric model that is able to separate out specific group-member influences. For this reason, I employ a set of spatial models to study the matter.

The randomness in the RPL allows for estimation of individual tastes. Individual-specific (conditional) estimates are obtained from the RPL estimates, assuming a triangular distribution for the random parameters. Specifically, individual-specific willingness to pay estimates are calculated by dividing the attribute estimate of interest by the marginal utility of income estimate. For example, the willingness to pay for Garmin trackers is found as:

$$WTP_i = -\frac{\beta_{S\_Garmin}}{\beta_{price,i}}$$

Subject-specific WTP for all attributes are then calculated in the same fashion and presented in Table 4.7.

In general, significant revisions in WTP are found in every attribute but Nike, with varying magnitudes. Prior to allowing for any peer influence, subjects have the highest WTP for GPS capacities, followed by watch design and sleep capacities. After learning about peer's experiences with activity trackers, and seeing their preferences, the sample subjects are significantly less willing to pay for the GPS and wristbands, while more willing to pay for all brand attributes. Among the other revisions, one notable result is that, before peer influence, the average WTP for wristband design is positive, ($129) whereas the WTP for the same design drops (by $152) below zero after peer influence. This shows that peer discussion plays a negative role with respect to wristband design. For example, subjects may have discussed the disadvantages of a design that prevent people from wanting to include this attribute.

I also expected asymmetric responses to peer recommendations: Peers may provide either positive or negative feedback, each with a different effect on changes in WTP. Brand proved to be a topic of much discussion among subjects. This discussion clearly had an impact on subjects' tendency to revise their valuations as prior to peer recommendations consumers' preferences depend on both brands and specific attributes such as design, function and price. After peer recommendations, however, subjects revise their preferences more positively on brands instead of specific attributes. Table 4.7 shows that peer recommendations positively enhanced brand knowledge. All brand attribute preferences (Garmin, Nike and Jawbone) are revised higher after peer influence compared to the baseline (Fitbit), with Garmin being revised higher by $95, Nike revised by $36; and Jawbone revised by $63.

Table 4.7 Subject-specific Marginal WTP by Attribute

	Pre-influence		Difference in WTP	
	Mean WTP	Std. err	Mean WTP	Std. err
Garmin	39.2900*	25.7862	94.6470*	23.4557
Nike	63.0608*	53.1249	36.2421	23.5743
Jawbone	-5.9432	46.0103	62.8538*	40.8843
Watch	155.6978*	134.9679	-79.4125*	51.6551
Wristband	129.1619*	80.9048	-151.9357*	98.8290
Sleep	149.5709*	94.4735	-137.622*	89.5187
GPS	159.9482*	123.1506	-175.8147*	114.3615
MSG	105.2684*	75.6572	-87.2695*	56.7658

Note: a single asterisk indicates significance at a 5% level.

Subjects also exhibited a willingness to change preferences for design. For example, the WTP for the watch-style attribute after peer interaction is revised to be lower by $79. In the same way, the WTP for the wristband style after peer recommendations is negative, meaning that subjects' WTP for the wristband attribute drop to a point that they are not willing to pay for a tracker with this style. In terms of functions, the WTP for all functions are revised lower after peer recommendations. WTP for the function of tracking sleeping patterns is revised by $138, the WTP for the GPS function is revised by $176 via peer recommendation, and the WTP for the messaging function is revised by $87. Clearly, peer communications discouraged subjects from paying for additional functions. In general, peer discussions lead subjects to be more brand-conscious, and discouraged subjects from paying for every additional function. Comparing the first and second stage RPL estimates and WTP revisions, however, does not address the issue of how the definition of space affects preference revision. That is, I do not include the spatial weight matrix directly in the RPL comparisons. For this purpose, I estimate the extent of preference revision as moderated by each subject's location in the social-spatial network in the next section.

4.6.3 Results of Spatial Models

Spatial models suffer from an embarrassment of riches in terms of the ways in which relationships among network members can be defined[25]. In this study, I focus on two different ways: tie strength and source expertise. In this section, I examine the nature of preference revisions under each definition. To test the second and third hypotheses developed above, namely which definition generates larger preference revisions, and how credibility is related to attribute preference revisions, I compare the estimate of ρ obtained from a model that uses tie strength as a measure of social proximity and the estimate of ρ from using source expertise as the measure of social proximity. In this way, I can examine which mechanism is more likely to have influenced subjects' preferences between the first and second rounds. In Table 4.10, I present two SAR-SEM models in which Model 1 uses the tie strength as spatial weight ($W_1 = W_2 = social\ closeness$);

[25] Prior research shows many possible ways to define social relationships, including frequency of communication (Goldenberg et al. 2009), degree of acquaintance (Godes and Mayzlin 2009), respect or leadership (Mullen, Johnson, and Salas 1991 ; Grippa and Gloor 2009), and trust (Buskens 1998 ; Berrera 2007). In this chapter I follow Richards, Hamilton and Allender (2014) in defining social closeness.

and Model 2 uses the source expertise as spatial weight ($W_1 = W_2 = perceived\ Credbility$).

The Lagrange-multiplier statistic tests the appropriateness of a model with a $\boldsymbol{W}$ matrix. In the case of tie strength, the LM value is 0.6919, which is smaller than the critical value of ${\chi^2}_1$ (with 1 degree of freedom), meaning that consumer preferences do not depend on peers that are defined through "social closeness". Moreover, unobserved correlation that cannot be explained with the social closeness matrix, as evident by the failure to reject the hull hypothesis of a spatial error specification (LM=1.2959<${\chi^2}_1$). On the other hand, the LM statistics for spatial lag (3.0414) and spatial error (3.2679) using peers defined by "source credibility" are both significant, rejecting the null hypothesis of an OLS in favor of both spatial lag and error specification. This means that source credibility is able to explain the causal relationship between consumer preferences and peers' preferences as well as the unobserved correlations. Based on these tests I conclude that consumers indeed depend on their peers (who they perceive as credible) for recommendation and that there exist unobserved correlations that cause consumers to arrive at the same choices.

Based on these results, the SAR-SEM model using tie strength does not reject the null hypothesis, my results do not support HYPOTHESIS 2 in that preference revisions will be greater for strong ties relative to weak ties. On the other hand, the SAR-SEM model with the "credibility" definition of social relationship rejects the null hypothesis of simple OLS in favor of the maintained SAR-SEM structure. This means that there is significant spatial dependence using reliability as weight matrix, providing supports for

HYPOTHESIS 3 in that perceived credibility is positively related to revisions in attribute preferences. The distinction is important, as it suggests that when facing innovative products, consumers do not turn to their friends for recommendations but rather people who they perceive as having expertise on the product. This extends the literature on proximity by identifying relational mechanisms through which social propinquity leads to information exchange.

Moreover, regarding weak ties those who are not social acquaintances, but possess perceived expertise on the subject of matter, are significant influencers. This finding is consistent with Granovetter (1993) in that "weak ties" rather than "friends" in a social network can be more influential. That is, people who have close social proximity are likely to share similar information, so they are not the best candidates for new product promotion. From a marketing perspective, this finding explains why online practices such as Yelp and TripAdvisor are successful, because they rely on expertise and experiences from strangers to promote their products and services. Since influential individuals are not necessarily close friends, this finding highlights a notable difference between traditional marketing procedures, where WOM plays an important role, and the more current, viral marketing where online recommendations are given by anyone who is perceived as credible. My finding is similar to Godes and Mayzlin (2009), who demonstrate that it is the less loyal customers instead of the most loyal customers who provide influential WOM. In a similar manner, my results suggest that marketers should get out of the traditional word-of-mouth marketing where friends recommend friends, and instead should target those who are credible representatives of the product.

The preferred model contained a number of other explanatory variables, but few showed a significant influence, independent of peer effects. The individual characteristics that were found to be significant in the preferred Model 2 were purchase frequency and BMI. This is intuitive as activity trackers are associated with fitness (within the healthy range of BMI) and the WTP is determined by how often an individual is likely to shop for sports goods. The sign for Purchase is positive, meaning that people who shop for sports goods more frequently are more willing to pay higher price for Jawbone. The sign for the BMI is also positive in Model 2, meaning that the higher a subject's BMI, the more likely he/she will choose Jawbone. This finding is intuitive as Jawbone is positioned as a lower-end product, with lower prices and fewer innovations. In other words, Jawbone is an introductory tracker, which is preferred by people who are new users, and perhaps not dedicated to physical fitness. Less fit subjects tend to have higher BMIs, hence the positive relation between Jawbone and BMI. The finding that BMI is an important factor in consumers' choices of activity tracker is not surprising as the main function of an activity tracker is to record physical activities. Besides individual-specific characteristics such as BMI, other factors that refer to more general characteristics of the sample are also apparent.

Two contextual effects were found to be significant: age and gender. Age is negatively related to the WTP for Jawbone, as subjects who are younger are likely to be more fitness aware. Gender is positively related, meaning that males are less likely to pay a higher price for a Jawbone. Note that this does not necessarily mean that males are more into fitness, and hence more willing to pay for activity trackers than females, but that males are more likely to pay for Jawbone. Income, on the other hand, does not show a significant relationship with consumers' preferences for activity trackers. This notion can be explained that for each brand it has a range of trackers that satisfy consumers of different income levels, therefore indicators such age and gender that identify with certain traits of an activity tracker turn out to be determining factors. Identifying the relationship between contextual factors and preference revisions is important because contextual factors help marketers target a group of people with similar background. For example, in this case, older female college students that are willing to pay more for Jawbone trackers.

In the previous section, I showed that consumer preferences are significantly altered in the second stage, which provides support for my first hypothesis in that peer recommendations will lead to preference revision for the recommended option if individuals regard the attribute in question to be salient to the choice decision. To tie such revisions more clearly to peer effects, I also estimate a group fixed effect. The group fixed effect examines whether there exists a significant difference in WTP between control groups and treatment groups, and is measured by the estimate of δ. Group fixed-effect was found in Model 2: subjects that were assigned to treatment groups revised their WTP for Jawbone by $23.84 after peer influence. This is evidence that WTP can be influenced by peer recommendations. Combined with the fact that those influencers are not necessarily friends, this finding suggests that intense promotion from people who have perceived expertise is effective. Moreover, the group fixed is not significant trough revision by source expertise but not by strong ties, showing that the strong ties have no influence on peer preferences. For innovative products, people who are considered to have expertise on the products serve as more influential agents to promote the products.

Table 4.8 Results of Spatial Models.

Explanatory variables	OLS	t ratio	Model 1 using Tie strength	z-value	Model 2 using Source expertise	z-value
Purchase	-0.0699	-1.93703	-0.0560	-1.4432	0.0308**	1.9893
Workout	8.7206	1.6808	5.7500	0.9729	7.6060	0.1301
Tracker	0.0611	0.0444	5.5335	0.3884	9.2621	0.4830
BMI	-1.1913	-1.3878	-1.3750	-1.5180	0.7832*	1.6384
Contextual effects						
Age	0.1628**	2.5931	0.0039	0.0495	-0.0496*	-2.323
Gender	-2.0359*	-1.3344	2.8750	1.1848	0.7100*	1.9359
Income	-0.0001	-1.6211	-0.0000	-0.712	-0.0000	0.6147
Endogenous effect						
ρ	N.A.	NA	0.8125	1.032	0.8505*	3.5559
Unobserved effect						
λ	NA	NA	0.6050	0.4367	0.3977	0.5035
Fixed Group effect						
Treatment	-35.9161	-1.4364	-5.00	0.1779	23.8401*	2.7842
Model Fit						
Log likelihood	-314.668					
				p-value		p-value
LM(lag)			0.6919	0.2549	3.0474*	0.0808
LM(error)			1.2959	0.5041	3.2679*	0.0706

* Significant at 90%
** Significant at 95

4.7 Conclusion

Relationships are important ways for consumers to acquire information, because the creation of knowledge is a social process. This is especially true when consumers are considering the purchase of innovative, new products, as little is known prior to the release of truly new products. Despite the importance of social interaction as a vehicle for knowledge acquisition and the extensive literature on peer effects, limited research has made an effort to investigate how peers influence the adoption of innovative products. This chapter offers evidence as to how consumer preferences are revised through peer recommendations in the context of activity trackers.

In this chapter, I use a two-stage experiment to examine preference revision via peer recommendations. I detect factors that are important in consumers' choices of activity trackers. Among which, brand is a general representation of specific attributes and consumer recognition embodied in activity trackers. I find that brand-related information such as design, function, and price are significant when consumers choose to buy activity trackers. However, brand serves as a representation of all traits when consumers revise their preference according to peer recommendations. That is, when consumers seek information from their peers, they tend to generalize certain attributes, or make the connection between brand and other people (peers)' the discussion of attributes. This finding shed lights on how marketers can best use peer networks to promote innovative new products. That is, instead of promoting specific attributes of an activity tracker, marketers should link the innovative feature to the brand in general. For example, Garmin is the top brand in GPS. When promoting Garmin activity trackers, instead of emphasizing on the perks of the GPS function itself, the marketer could link Garmin activity trackers with excellent GPS performance compared to other brands.

Identifying the effect of peer relationships on consumer choice is a matter of both experimental design, and econometric estimation. In this study, my experiment is random in the sense students who are sampled are based on their preferences for academic majors, which should not correlate with their preferences for activity trackers. My econometric model is non-linear while addressing for two different mechanisms of peer recommendations. Peer recommendations work through the social proximity among members of the network. Such interaction is spatial and simultaneous in nature, which calls for spatial models. Spatial models allows for peer recommendations to enter through a weight matrix that address interrelationships, and yield true peer effect as a result.

I find that source credibility is more important in moderating social learning than social proximity. This provides evidence that individuals who are perceived to have expertise on the product rather than those they are friends with. Consistent with Granovetter's (1973) expectation that weak ties exhibit stronger interpersonal effects than do strong ties, I find that consumers do not revise their preferences according to how well they know each other, but rather how reliable they perceive their peers to be. Because my research products are activity trackers, people who are perceived to be reliable are those who dress in gym gear, are physically fit, and have previous experiences with activity trackers. These people serve as "hubs" in the network as they are the influencers of consumers' preference revision. In the broader sense of marketing, individuals who are perceived to have professional and reliable opinions of the subject of matter should introduce new products, rather than close friends and family.

Although this study is conducted in the context of activity trackers, my approach is applicable when studying other products that are innovative in nature and can be recommended via source credibility. Future research can extend this study in a number of ways: first, I did not consider information externalities, which could be another application of the data. Also, I used a choice-based conjoint experiment, which provided many observations from the same individual but suffers from the fact that attribute values do not vary over time. Replication with different items that vary in terms of their attribute content would help identify the model from this perspective.

CHAPTER 5. CONCLUSIONS AND IMPLICATIONS

Innovation is critically important to the agricultural sector. When new ideas are invented, diffused, and adopted, the benefits of superior products flow to either consumers in terms of preferred new attributes, or producers in terms of lower costs of production. Advantageous innovations, however, do not sell themselves, as there are many barriers on the path of diffusion from the lab to the ultimate end-user. Therefore, it is critical that we understand the nature of the diffusion process in order to ensure that it operates effectively, and efficiently. In this dissertation, I analyze innovation from the initial licensing stage to diffusion among farmers, to consumer preferences in the final market.

In the first essay, I study the optimal licensing strategies for university-based innovations. Passage of the Bayh-Dole Act allowed universities to license federally funded inventions created by their faculty, which enabled them to generate much-needed revenue from their labs. Doing so, however, raises both economic and legal issues as to whether a university should limit exclusive access to the government-funded innovations? Moreover, should patents be licensed using a system of royalties, or with fixed licensing fees?

I address the economic dimension of these questions using an oligopolistic model of downstream competition in which two companies compete in price for a quality improving innovation. With this model, I argue that the most salient feature of the demand for new products is the fundamental asymmetry in the demand for improved relative to existing products. I show that non-exclusive licensing provides the maximum revenue to the university, as well as the downstream firms, by aligning the benefits of both. Consumers have differentiated tastes and are willing to pay more for high quality products. By licensing non-exclusively, the patent holder "wipes out" low quality products, but is able to compensate for the loss of demand with higher margins. Moreover, because the consumer market consists of only high quality products, university researchers have an incentive to take risks and create bold new products because the premium for higher quality is sufficient to compensate for R&D investment. Ultimately, my finding imply that Bayh-Dole Act is likely welfare-improving as consumers are left with a market uniformly supplied with high-quality products, all firms generate higher profits, and universities have a lucrative source of funding.

Preference for one method of licensing over another is driven entirely by differences in their impact on strategic behavior downstream. Namely, I find that if downstream firms pay fixed fees, competition is softened between the two competing downstream firms. On the other hand, a system of royalties intensifies downstream competition. Fixed fees provide the technology transfer office (TTO) an opportunity to extract a lump sum of money up front, but yield no control over the output in the consumer market. Royalties, meanwhile, become part of marginal costs, and therefore increase with the amount of output. In this way, royalties allow the TTO to control demand in the final market, and enhance competition.

My findings have practical implications, both for settled controversies regarding university licensing, and potential future challenges to the Bayh-Dole Act. In one prominent example I describe in Chapter 2, the "SweeTango Case", the essential legal question concerns whether growers should be denied access to the research products (namely, the SweeTango apples) developed by their own land-grant university? Prior to developing the SweeTango, the owner of the patent, the University of Minnesota, earned royalties of $8 million per year through another popular apple variety: the Honeycrisp.[26] By licensing exclusively to a single group of growers, the university was trying to achieve similar success with SweeTango. However, my results show that it is in the best interest of the university to license non-exclusively rather than limiting the access to its new variety. Not only does non-exclusive licensing maximizes social welfare – consistent with the arguments of the plaintiff in this case, the growers denied access to SweeTango -- but it generates the most profit for the university. Patent licensing, in general, allows universities, researchers, and downstream firms all to benefit from innovation. A robust licensing strategy provides incentives for universities to conduct marketable research, and commercialize new products for valuable end-use, which is exactly the purpose of the Bayh-Dole Act.

After they are licensed, innovations must be adopted to be of economic importance. Observed adoption patterns exhibit two curious patterns: First, even clearly viable innovations are typically only adopted after a substantial lag and, second, adoption in a geographic region tends to show a marked "clustering" pattern, or nodes of adopters that are far from randomly distributed. In this dissertation, I explain each of these patterns as a result not of a market failure, but rather as a rational response to economic incentives.

Particularly in developing countries, adoption occurs only with a considerable lag. Previous research identifies a lack of information as the primary barrier to rapid adoption

[26] Read the news at: http://news.yahoo.com/apnewsbreak-settlement-sweetango-apple-lawsuit-203025186.html

of beneficial inventions. In developing countries, farmers often have insufficient information due to limited access to extension services, leading to low adoption rates and delays in adoption. In the absence of formal extension services, farmers typically learn from their neighbors. Information provided by early adopters may either speed up or slow down the diffusion process. When there are no or few early adopters, farmers may delay adoption in anticipation of information from others' experiences, which is termed a "strategic delay" as the farmer waits to adopt as a rational strategy to maximize the benefit of new information.

Learning from neighbors is also one plausible explanation for the observed clusters of adopters and non-adopters in my sample of Mozambique maize farmers. Learning from neighbors requires that one farmer first experiments with the innovation, has a good or bad experience, and then others draw their own conclusions based on their observations, or conversations with the adopter. Learning from others normally happens with one or few "leaders" adopt the new variety. After the early adopters demonstrate the success of the variety, others will follow on the assumption that their experience will be similar. Learning from neighbors may be as simple as mimicry, or it may be true learning or other adaptive behavior. Either way, clusters of adopters form.

I formally test the importance of learning using a farm-level data set of maize variety adoption in Mozambique. My econometric model allows for the spatial nature of social interactions. In particular, the model captures several potential forms of interpersonal relationship among farmers, and their impact on variety adoption. I find that clusters are indeed formed by farmers following each other's decisions, presumably through either direct communication, observation of some other indirect form of imitation. When a farmer adopts the new variety, others imperfectly observe how well it performs. Depending on how closely farmers communicate, and the nature of the new variety, non-adopters may derive differing degrees of information about the performance of the new variety. They then decide whether it is likely to be profitable for the next planting season. When there are few adopters in a network, farmers will strategically postpone adoption in order to acquire more information, leading to clusters of non-adopters.

Despite the fact that direct communication and indirect observation regarding new varieties help convey information about their likely attractiveness, which helps speed adoption, gaining information from others can also lead farmers to delay adoption in anticipation of more information. If few farmers have adopted, the delay effect dominates and more social interaction can actually slow adoption. Such strategic delay is well documented in other contexts, but mine is the first evidence in an explicitly spatial model of social interaction. A spatial model is necessary to test the strategic delay hypothesis because the underlying mechanism is implicitly spatial: Communication and observation take place between neighbors much more readily than between those who are distant.

On a deeper level, I find that clustering is driven not only by the geographic proximity of neighbors, but the specific structure of the social-spatial network. I measure the structure of a farmer's network in three ways, resulting in three different spatial weight matrices: (1) a nearest-neighbor matrix, in which the elements assume values of 1 for farmers who are nearest neighbors to each other, and 0 otherwise, (2) a four-neighbor extension in which the spatial matrix is populated by 1s only for those who share boundaries, and (3) a geographic distance in which the distance between pairs of farmers is measured using Euclidean space. Geographical distance is found to be insignificant, so I conclude that a simplistic measure of proximity is not an appropriate measure of network structure. Instead, farmers depend on their immediate neighbors and extended neighbors (who share boundaries) for information that leads to adoption decisions.. Immediate and extended neighbors serve as the most accessible information source in developing countries due to imitated access to official extension services. However, learning from neighbors is not constrained by geographic distance because farmers go beyond immediate neighbors for new information. Clearly, learning is critical to adoption. In fact, I find that a number of other sources of information are statistically important. For example, I find that education and extension programs are drivers of new variety adoption. When provided, extension increases farmers' knowledge of the unfamiliar variety, therefore, reduces the risks of adopting. Economies of scale also proxy the ability to obtain more, and better, information. Because information-acquisition is largely a fixed cost, larger farmers have a greater incentive to obtain more information on a cost-benefit basis. Scale also helps to reduce the risk of adopting because larger farms produce output for commercial sale rather than home-consumption, so have a more robust supply-chain.

Once new products are licensed, and adopted by producers, their ultimate success is determined in the end-consumer market. New products, especially expensive and innovative products such as fitness trackers, represent a risky purchase to consumers who are not familiar with the relative merits of various attributes. Therefore, consumers typically rely on peers for help in decision-making. However, true peer effects – the direct effect of agents changing their behavior after observing from, learning from, or mimicking others -- are difficult to identify econometrically because they represent only one reason why people in a social network tend to behave similarly.

The empirical example in the third essay helps address the identification problem, both conceptually and empirically. First, I show how random assignment in a lab experiment can be used to identify true peer effects. By introducing peer information into only the treatment groups, I find significant changes in the willingness to pay regarding specific tracker attributes. Second, I show how spatial models can help identify peer effects. Identification requires variation in peer-exposure across individual subjects. Spatial models generate the necessary variation because spatial-weight matrices describe each pairwise relationship between sample members. Third, I exploit the richness of spatial econometrics to test the relative importance of different types of relationships among consumers on their adoption decisions. Spatial-weight matrices not only capture each pairwise relationship, but can describe different types of relationship within dyads, whether two subjects are friends, one is expert and the other uninformed, or one is passive and the other aggressive.

Specifically, I differentiate between alternative definitions of space in order to identify the precise mechanism through which peer effects operate. Namely, I introduce tie strength and source credibility in order to measure the importance of different types of peer influence. Traditional marketing focuses on word-of-mouth among family and friends (strong ties). However, I find that people that are perceived as having credible opinions of trackers serve as the opinion leaders. In other words, when consumers face the choice of an innovative product, they are more likely to adopt opinions from those who are "experts" in this field rather than people with whom they maintain close relationships. For example, Yelp and TripAdvisor are examples of business models that utilize source expertise. With either, consumers typically rely on reviews from others when they are looking for a new product or service. Anyone can provide a review, however, those that are considered more credible are more trusted, and more effective. Online marketing tools rely on viral marketing, which relaxes the constraints faced by traditional word-of-mouth in providing a more efficient means of exchanging information.

My findings form an important contribution to the existing literature by addressing the dynamic aspect of each stage of innovation diffusion. Although innovation consists of three independent operations: the invention stage, the initial distribution of new product, and the mass adoption by consumers, I address all three stages with a focus on the interactive nature among players. Licensing consists of a game between upstream patent holder and the downstream firms, and is only socially optimal when the patent holder recognizes the strategic interdependence of downstream buyers. Adoption of an innovation by growers, or manufacturers more generally, relies on experience and learning from others, which is also highlights the interdependence of otherwise-independent agents. Finally, diffusion among downstream consumers also involves interaction among end-users who differ in terms of the amount of information they may possess, their intensity of demand for the product, and their inherent willingness to adopt. Therefore, although each stage of the process may be functionally independent, the processes involved are similar, and should be recognized as such.

My findings go beyond agriculture, and likely apply to product development and diffusion in many other contexts. For example, a company called Activate Networks applies technology license from Harvard University to draw connections between physicians with patients in common, then uses those ties to accelerate the adoption of newly introduced drugs. In terms of diffusion, they are seeking to identify the credible opinion leaders, then use the network effect to amplify the company's own promotional performances. In fact, the healthcare industry is one in which innovation is a constant, and yet adoption is seldom immediate. Responding to the high demand for the commercialization of university-sourced inventions, technology transfer offices and third-party organizations have taken on the role of licensing university-based innovations. Many major universities such as Johns Hopkins, University of Maryland, and Arizona State University have established separate departments solely for the purpose of technology commercialization. Other "outside" technology transfer organization such as BioHealth Innovation also serve to evaluate market-relevant bio heath intellectual properties (IP), connect the IP with funding, and assist in marketing. My research provides a guideline for these offices and organizations that are looking to optimize university profit, all while maximizing social welfare derived by new drugs.

Similar to the adoption of agricultural innovations, the diffusion of new drugs among consumers often focuses on interpersonal communication within a social group. In promoting new drugs, direct marketing is not the optimal marketing strategy for two reasons: (1) there is a limitation to how successful a pharmaceutical company can be in encouraging members of a social group (for example, doctors) to adopt a new product or practice; and (2) there is no way that they can be seen as credible by everyone or address the individual concerns of each person. My research shows that consumers depend on opinion leaders who are perceived as credible, which suggests that alterative marketing strategies such as WebMD and ASK.COM as potential ventures. Like Yelp, WebMD is a consumer-rated website that offers professional medical opinions, with its usefulness rated by users. After the discussion of each symptom, a recommended treatment is provided. When consumers rate particular recommendations as more credible and useful, new users are more likely to adopt the same medication. At the same time, they are more likely to leave positive feedback so that more people will adopt. Indeed, bringing network analysis to the study of diffusion processes addresses the observation that if you can persuade someone to adopt a new product, and then promote it to others, they will have a global impact on adoption within the group a whole.

Each stage of the research presented here is not without limitations. In the first essay, my results regarding licensing are based on the assumed duopoly-Bertrand framework. If conduct in the downstream market is something other than Bertrand, I may have found different conclusions. Future research may want to increase the number of downstream firms under other oligopolistic scenarios, or even perfect competition. In the second essay, although Mozambique is representative of the adoption case in developing countries, there may be differences between cash crops and staple foods produced in other countries. Cash crops are planted for profitability rather than own-consumption, so the adoption decision may involve considerations other than the ones discussed here. In the third essay, my consumer preference experiment can be extended to analyze other innovative products. Future research may use an auction instead of a choice experiment, as auctions may be more effective in eliciting true willingness to pay relative to a hypothetical choice experiment.

REFERENCES

Abdulai, Awudu, and Wallace E. Huffman. 2005. The diffusion of new agricultural technologies: The case of crossbred-cow technology in Tanzania. *American Journal of Agricultural Economics* 87, no. 3:645-659.

Anderson, Simon P., De Palma, Andre, and Thisse, Jacques F. 1992. *Discrete choice theory of product differentiation.*MIT press.

Anselin, Luc. 2003. Spatial externalities, spatial multipliers, and spatial econometrics. *International regional science review* 26, no. 2:153-166.

Anselin, Luc. 2002. Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics* 27, no. 3:247-267.

Bala, Venkatesh, and Sanjeev Goyal. 2000. A noncooperative model of network formation. *Econometrica* 68, no. 5:1181-1229.

Bala, Venkatesh, and Sanjeev Goyal. 2000. A noncooperative model of network formation. *Econometrica* 68, no. 5:1181-1229.

Bandiera, Oriana, and Imran Rasul. 2006. Social networks and technology adoption in northern mozambique*. *The Economic Journal* 116, no. 514:869-902.

Banerjee, Abhijit V. 1992. A simple model of herd behavior. *The Quarterly Journal of Economics*797-817.

Banerjee, Abhijit V. 1992. A simple model of herd behavior. *The Quarterly Journal of Economics*797-817.

Ben-Akiva, Moshe E., and Lerman, Steven R. 1985. *Discrete choice analysis: theory and application to travel demand.* Vol. 9. MIT press.

Berndt, Ernst R., Robert S. Pindyck, and Pierre Azoulay. 1999. *Network effects and diffusion in pharmaceutical markets: antiulcer drugs*.

Besley, Timothy, and Anne Case. 1994. *Diffusion as a learning process: Evidence from HYV cotton*.

Besley, Timothy, and Anne Case. 1993. *Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits*.

Besley, Timothy, and Stephen Coate. 1997. An economic model of representative democracy. *The Quarterly Journal of Economics*85-114.

Bias, Calisto, and Cynthia Donovan. 2003. Gaps and opportunities for agricultural sector development in Mozambique.

Bousquet, Alain, Helmuth Cremer, Marc Ivaldi, and Michel Wolkowicz. 1998. Risk sharing in licensing. *International Journal of Industrial Organization* 16, no. 5:535-554.

Brock, William A., and Steven N. Durlauf. 2001. Discrete choice with social interactions. *The Review of Economic Studies* 68, no. 2:235-260.

Cahoon, RS, et al. 2007. Licensing agreements in agricultural biotechnology. *Intellectual property management in health and agricultural innovation: a handbook of best practices, Volumes 1 and 2*1009-1016.

Caplin, Andrew, and John Leahy. 1998. Miracle on sixth avenue: information externalities and search. *The Economic Journal* 108, no. 446:60-74.

Carney, Monica H. 2013. Identifying Peer Effects: Thinking outside the 'Linear-in-Means' Box.

Case, Anne. 1992. Neighborhood influence and technological change. *Regional Science and Urban Economics* 22, no. 3:491-508.

Case, Anne C. 1991. Spatial patterns in household demand. *Econometrica: Journal of the Econometric Society*953-965.

Chamley, Christophe, and Douglas Gale. 1994. Information revelation and strategic delay in a model of investment. *Econometrica: Journal of the Econometric Society*1065-1085.

Chamley, Christophe, and Douglas Gale. 1994. Information revelation and strategic delay in a model of investment. *Econometrica: Journal of the Econometric Society*1065-1085.

Chamley, Christophe, and Douglas Gale. 1994. Information revelation and strategic delay in a model of investment. *Econometrica: Journal of the Econometric Society*1065-1085.

Choi, Jay P. 2001. Technology transfer with moral hazard. *International Journal of Industrial Organization* 19, no. 1:249-266.

Choi, Jay P. 1997. Herd behavior, the" penguin effect," and the suppression of informational diffusion: an analysis of informational externalities and payoff interdependency. *The Rand journal of economics*407-425.

De Giorgi, Giacomo, Michele Pellizzari, and Silvia Redaelli. 2010. Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*241-275.

Dolan, Brian. 2014. *Fitbit, Jawbone, Nike had 97 percent of fitness tracker retail sales in 2013.*

Duflo, Esther, and Emmanuel Saez. 2003. *Implications of information and social interactions for retirement saving decisions*.

Duflo, Esther, and Emmanuel Saez. 2002. *The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment*.

Ellison, Glenn, and Drew Fudenberg. 1995. Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*93-125.

Ellison, Glenn, and Drew Fudenberg. 1993. Rules of thumb for social learning. *Journal of Political Economy*612-643.

Elrod, Terry, Jordan J. Louviere, and Krishnakumar S. Davey. 1992. An empirical comparison of ratings-based and choice-based conjoint models. *Journal of Marketing Research* 29, no. 3:368-377.

Fauli-Oller, Ramon, and Joel Sandonis. 2002. Welfare reducing licensing. *Games and Economic Behavior* 41, no. 2:192-205.

Feder, Gershon, and Gerald T. O'Mara. 1982. On information and innovation diffusion: A Bayesian approach. *American Journal of Agricultural Economics* 64, no. 1:145-147.

Fishbein, Martin E. 1967. Readings in attitude theory and measurement. .

Foster, Andrew D., and Mark R. Rosenzweig. 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*1176-1209.

Gallini, Nancy T., and Brian D. Wright. 1990. Technology transfer under asymmetric information. *The Rand journal of economics*147-160.

Ghadim, Amir K. A., and David J. Pannell. 1999. A conceptual framework of adoption of an agricultural innovation. *Agricultural Economics* 21, no. 2:145-154.

Ghadim, Amir K. A., and David J. Pannell. 1999. A conceptual framework of adoption of an agricultural innovation. *Agricultural Economics* 21, no. 2:145-154.

Goolsbee, Austan, and Peter J. Klenow. 1999. *Evidence on learning and network externalities in the diffusion of home computers*.

Griliches, Zvi. 1957. Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*501-522.

Griliches, Z. 1960. Hybrid Corn and the Economics of Innovation. *Science (New York, N.Y.)* 132, no. 3422:275-280.

Hensher, David A., Rose, John M., and Greene, William H. 2005. *Applied choice analysis: a primer.*Cambridge University Press.

Jensen, Richard. 1982. Adoption and diffusion of an innovation of uncertain profitability. *Journal of Economic Theory* 27, no. 1:182-193.

Jovanovic, Boyan, and Yaw Nyarko. 1995. A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity.Microeconomics*247-305.

Kamien, Morton I., Shmuel S. Oren, and Yair Tauman. 1992. Optimal licensing of cost-reducing innovation. *Journal of Mathematical Economics* 21, no. 5:483-508.

Kamien, Morton I., and Yair Tauman. 2002. Patent licensing: the inside story. *The Manchester School* 70, no. 1:7-15.

Kamien, Morton I., and Yair Tauman. 1986. Fees versus royalties and the private value of a patent. *The Quarterly Journal of Economics*471-491.

Kamien, Morton I., and Tauman, Yair. 1984. *The private value of a patent: a game theoretic analysis.*Springer.

Kapur, Sandeep. 1995. Technological diffusion with social learning. *Journal of Industrial Economics* 43, no. 2:173-195.

Katz, Michael L., and Carl Shapiro. 1985. Network externalities, competition, and compatibility. *The American Economic Review*424-440.

Kelejian, HH. und DP Robinson (1998)„„Spatial Correlation: A Suggested Alternative to the Autoregressive Model. *Anselin/Florax*75-95.

Kuhn, Peter, Peter Kooreman, Adriaan Soetevent, and Arie Kapteyn. 2011. The effects of lottery prizes on winners and their neighbors: Evidence from the Dutch Postcode Lottery. *The American Economic Review*2226-2247.

Leathers, Howard D., and Melinda Smale. 1991. A Bayesian approach to explaining sequential adoption of components of a technological package. *American Journal of Agricultural Economics* 73, no. 3:734-742.

Lee, Lung-fei. 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics* 140, no. 2:333-374.

Lee, Lung- Fei. 2004. Asymptotic Distributions of Quasi- Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica* 72, no. 6:1899-1925.

Li, Changying, and Junmei Wang. 2010. Licensing a Vertical Product Innovation*. *Economic Record* 86, no. 275:517-527.

Lin, Xu. 2010. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics* 28, no. 4:825-860.

Lopes, Helder. 2010. *Adoption of improved maize and common bean varieties in Mozambique*.

Louviere, Jordan J., Hensher, David A., and Swait, Joffre D. 2000. *Stated choice methods: analysis and applications.*Cambridge University Press.

Louviere, Jordan J., and George Woodworth. 1983. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research*350-367.

Mansfield, Edwin. 1961. Technical change and the rate of imitation. *Econometrica: Journal of the Econometric Society*741-766.

Manski, Charles F. 2000. *Economic analysis of social interactions*.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *The review of economic studies* 60, no. 3:531-542.

McFadden, Daniel. 1974. The measurement of urban travel demand. *Journal of public economics* 3, no. 4:303-328.

McFadden, Daniel L., and Kenneth E. Train. 1996. Consumers' evaluation of new products: Learning from self and others. *Journal of Political Economy*683-703.

Moffitt, Robert A. 2001. Policy interventions, low-level equilibria, and social interactions. *Social dynamics* 4, 45-82.

Motta, Massimo. 1993. Endogenous quality choice: price vs. quantity competition. *The journal of industrial economics*113-131.

Narayan, Vishal, Vithala R. Rao, and Carolyne Saunders. 2011. How peer influence affects attribute preferences: a Bayesian updating mechanism. *Marketing Science* 30, no. 2:368-384.

Oh, Kyoungwhan. 2014. The Effects of Brand, Design, and Price on Intent to Purchase an Activity Tracker.

Prax, Valerie. 2010. *American Family Farmers Feed 155 People Each- 2% Americans Farm.*

Rao, Vithala R. 2014. *Applied conjoint analysis.*Springer.

Richards, Timothy J., Stephen F. Hamilton, and William J. Allender. 2014. Social networks and new product choice. *American Journal of Agricultural Economics*aat116.

Rosenzweig, Mark R. 1995. Why are there returns to schooling? *The American Economic Review*153-158.

Roumasset, James A. 1976. *Rice and risk. Decision making among low-income farmers.*North Holland Publ. Comp.

Saracho, Ana I. 2002. Patent licensing under strategic delegation. *Journal of Economics & Management Strategy* 11, no. 2:225-251.

Sen, Debapriya. 2005. Fee versus royalty reconsidered. *Games and Economic Behavior* 53, no. 1:141-147.

Sen, Debapriya, and Yair Tauman. 2007. General licensing schemes for a cost-reducing innovation. *Games and Economic Behavior* 59, no. 1:163-186.

Sen, Debapriya, and Yair Tauman. 2007. General licensing schemes for a cost-reducing innovation. *Games and Economic Behavior* 59, no. 1:163-186.

Shapiro, Carl. 1985. Patent licensing and R & D rivalry. *The American Economic Review*25-30.

Singh, Nirvikar, and Xavier Vives. 1984. Price and quantity competition in a differentiated duopoly. *The Rand journal of economics*546-554.

Soetevent, Adriaan R. 2006. Empirics of the Identification of Social Interactions; An Evaluation of the Approaches and Their Results*. *Journal of Economic Surveys* 20, no. 2:193-228.

Stamatopoulos, Giorgos, and Yair Tauman. 2008. Licensing of a quality-improving innovation. *Mathematical Social Sciences* 56, no. 3:410-438.

Statt, Nick. 2014. *Nike FuelBand -- out of gas already?*.

Sunding, David, and David Zilberman. 2001. The agricultural innovation process: research and technology adoption in a changing agricultural sector. *Handbook of agricultural economics* 1, 207-261.

Thurstone, Louis L. 1927. A law of comparative judgment. *Psychological review* 34, no. 4:273.

Train, Kenneth, and Weeks, Melvyn. 2005. *Discrete choice models in preference space and willingness-to-pay space.*Springer.

Uaiene, Rafael N., C. Arndt, and WA Masters. 2009. Determinants of agricultural technology adoption in Mozambique. *Discussion papers*no. 67E:1-29.

Wang, X. H. 2002. Fee versus royalty licensing in a differentiated Cournot duopoly. *Journal of economics and business* 54, no. 2:253-266.

Warner, Kenneth E. 1974. The need for some innovative concepts of innovation: an examination of research on the diffusion of innovations. *Policy Sciences* 5, no. 4:433-451.

Wilson, Robert. 1975. Informational economies of scale. *The Bell Journal of Economics*184-195.

Yang, Sha, and Greg M. Allenby. 2003. Modeling interdependent consumer preferences. *Journal of Marketing Research* 40, no. 3:282-294.

Zhang, Jianbo. 1997. Strategic delay and the onset of investment cascades. *The Rand journal of economics*188-205.

Zhao, Jinhua. 2003. *Information externalities and strategic delay in technology adoption and diffusion*.

Zhao, Jinhua. 2003. *Information externalities and strategic delay in technology adoption and diffusion*.

Zimmerman, David J. 2003. Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and Statistics* 85, no. 1:9-23.

APPENDIX A

CACULATING INNOVATOR'S PROFIT

In this appendix, I derive the expressions for the inverse demand functions, first- and second-order derivatives for the innovator's profit function, optimal royalties, and innovator profits under exclusive royalty, and two-part tariff licensing schemes.

To get the inverse demand function for firm i, we take derivative of the utility function with respect to q_i.

$$p_1 = \frac{dU(q_1,q_2)}{dq_1} = c(s_1) - q_1 - bq_2,$$

for inverse demand of firm 1, and:

$$p_2 = \frac{dU(q_1,q_2)}{dq_2} = c(s_2) - q_2 - \lambda bq_1$$

or inverse demand of firm 2.

To get the expression of the optimal prices for firm 1 and 2, we take derivative of the prices with respect to each profit function and set them to 0.

$$\frac{\partial \pi_2^{RE}}{\partial p_2} = \frac{1 - 4p_2\lambda^2 + 2r\lambda^2 - b\lambda^3 + 2bp_1\lambda^3}{2\lambda^2 - 2b^2\lambda^3} = 0$$

$$\frac{\partial \pi_1^{RE}}{\partial p_1} = \frac{b + (-1 + 4p_1)\lambda^2 - 2bp_2\lambda^2}{2\lambda^2(-1 + b^2\lambda)} = 0$$

Solving the above functions together, optimal price for firm 2 can be expressed as a function of b, r and λ:

$$p_2 = -\frac{2 - b^2\lambda + 4r\lambda^2 - b\lambda^3}{2\lambda^2(-4 + b^2\lambda)}$$

Substitute the above expression of p_2 into the demand function to get q_2, then we get the expression of the innovator's profit as a function of b, r and λ.

$$\pi_3^{RE} = rq_2 - c_3 = -\frac{-2r^2\lambda^2(-2+b^2\lambda) + 2(4-5b^2\lambda+b^4\lambda^2) + r(-2+b^2\lambda+b\lambda^3)}{2\lambda^2(4-5b^2\lambda+b^4\lambda^2)}$$

Solving for the first order condition of the above function with respect to r and set it to 0:

$$\frac{\partial \pi_3^{RE}}{\partial r} = \frac{2-8r\lambda^2-b\lambda^3+b^2\lambda(-1+4r\lambda^2)}{2\lambda^2(4-5b^2\lambda+b^4\lambda^2)} = 0$$

I get the optimal r

$$r^{RE} = \frac{-2+b^2\lambda+b\lambda^3}{4\lambda^2(-2+b^2\lambda)}$$

Substitute the expression for r back to the profit function π_3^{RE} to get the profit as a function of sole b and λ.

$$\pi_3^{RE} = \frac{-16b^6\lambda^5 + b^4\lambda^2(112\lambda^2-1) - 2b^3\lambda^4 - b^2\lambda(\lambda^5+224\lambda^2-4) + 4b\lambda^3 + 128\lambda^2 - 4}{16\lambda^4(b^2\lambda-4)(b^2\lambda-2)(b^2\lambda-1)}$$

After solving for the optimal fee and royalty, the innovator's profit under exclusive two-part tariff and non-exclusive two-part tariff licensing, both as a function of b and λ are given by:

$$\pi_3^{TE} = F^{TE} + r^{TE}q_2^{TE} - \frac{1}{\lambda^2}$$
$$= \frac{-4b^9\lambda^5 + 12b^8\lambda^5 + b^7\lambda^2(\lambda^5+36\lambda^2-1)}{(b-2)^2(b+1)\lambda^4(b^2\lambda-6)(b^2\lambda-2)(b^2\lambda-1)} +$$
$$\frac{-b^6\lambda^2(\lambda^5+16\lambda^3+110\lambda^2-3) + b^5(-10\lambda^6+6\lambda^4-80\lambda^3+4\lambda)}{4(b-2)^2(b+1)\lambda^4(b^2\lambda-6)(b^2\lambda-2)(b^2\lambda-1)} +$$
$$\frac{4b^4\lambda(3\lambda^5+36\lambda^3+61\lambda^2-\lambda-3) + 4b^3(5\lambda^5-2\lambda^4-3\lambda^3+12\lambda^2-1)}{(b-2)^2(b+1)\lambda^4(b^2\lambda-6)(b^2\lambda-2)(b^2\lambda-1)} +$$
$$\frac{-4b^2(\lambda^6+5\lambda^5+80\lambda^3+36\lambda^2-4\lambda-3) - 4b\lambda^3(3\lambda-4) + 4(3\lambda^4+48\lambda^2-4)}{(b-2)^2(b+1)\lambda^4(b^2\lambda-6)(b^2\lambda-2)(b^2\lambda-1)}$$

for profit under exclusive two-part tariff licensing.

$$\pi_3^{TN} = 2F^{TN} + r^{TN}(q_1^{TN} + q_2^{TN}) - \frac{1}{\lambda^2} =$$

$$\frac{2b^{13}\lambda^8+2b^{12}\lambda^7(3\lambda+2)+b^{11}\lambda^4(\lambda^6+18\lambda^3+6\lambda^2-1)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{-b^{10}\lambda^4(\lambda^6+2\lambda^5+8\lambda^4+90\lambda^3+78\lambda^2-3)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{b^9\lambda^3(-13\lambda^6-3\lambda^5+16\lambda^4+32\lambda^3-90\lambda^2+2\lambda+15)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{b^8\lambda^3(15\lambda^6+33\lambda^5+120\lambda^4+472\lambda^3+566\lambda^2-10\lambda-49)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{b^7\lambda^2(46\lambda^6+45\lambda^5-252\lambda^4-632\lambda^3+444\lambda^2+4\lambda-80)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{-b^6\lambda^2(80\lambda^6+197\lambda^5+592\lambda^4+1136\lambda^3+1864\lambda^2-84\lambda-256)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{4b^5\lambda(7\lambda^6-53\lambda^5+302\lambda^4+373\lambda^3-202\lambda^2-10\lambda+39)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{4b^4\lambda(2\lambda^7+33\lambda^6+126\lambda^5+256\lambda^4+540\lambda^3+664\lambda^2-86\lambda-121)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{-4b^3(4\lambda^7+48\lambda^6-91\lambda^5+540\lambda^4+234\lambda^3-124\lambda^2-12\lambda+24)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{-4b^2(24\lambda^6+135\lambda^5+128\lambda^4+792\lambda^3+316\lambda^2-156\lambda-72)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}+$$
$$\frac{8b\lambda(18\lambda^4-27\lambda^3+156\lambda^2-8)+24(9\lambda^4+72\lambda^2-16)}{2(b-2)^2(b+1)\lambda^4(b\lambda-3)(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)}$$

for profit under non-exclusive two-part tariff licensing.

I also derive expressions for downstream profits under exclusive and non-exclusive two-part tariff licensing. In each case, I show that the resulting maximum profit is less than that available in the benchmark, Nash sub-game equilibrium in the second-stage of the licensing game, so licensing will not occur. Maximizing innovator profit with respect to the royalty and the fixed fee, and substituting the resulting expressions back into the downstream profit functions leads to equilibrium profits for each firm. In the exclusive licensing case, the profit for firm 2 is found from the equilibrium as the second-stage game to be:

$$\pi_2^{TE} = (p_2^{TE} - r^{TE})q_2^{TE} - F^{TE} =$$
$$\frac{b^9(-\lambda^8) + b^8\lambda^8 + b^7\lambda^2(15\lambda^5 - 2) + b^6(-15\lambda^7 - 4\lambda^4 + 6\lambda^2)}{4(b-2)^2(b+1)\lambda^4(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$
$$\frac{b^5(-76\lambda^6 + 12\lambda^4 + 8\lambda) + 8b^4\lambda(10\lambda^5 + \lambda^2 - \lambda - 3)}{4(b-2)^2(b+1)\lambda^4(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$
$$\frac{4b^3(33\lambda^5 - 4\lambda^4 - 6\lambda^3 - 2) - 4b^2(2\lambda^6 + 33\lambda^5 - 8\lambda - 6)}{4(b-2)^2(b+1)\lambda^4(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$
$$\frac{8b\lambda^3(9\lambda - 4) + 8(9\lambda - 4)}{4(b-2)^2(b+1)\lambda^4(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} < \pi^{NL}$$

while firm 1 does not license by the definition of the exclusive contract.

In the non-exclusive case, the maximum profit for both firms under two-part tariff licensing is:

$$\pi_1^{TN} = \pi_2^{TN}$$

$$= \frac{-b^{12}\lambda^{11} + 24b\lambda(9\lambda^4 - 27\lambda^3 + 12\lambda^2 - 4)) + 72(9\lambda^4 - 8)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{b^2(-432\lambda^6 - 1404^5 + 96\lambda^4 + 160\lambda^2 + 816\lambda + 432)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{b^6\lambda^2(-268\lambda^6 - 379\lambda^5 - 24\lambda^4 - 4\lambda^3 + 4\lambda^2 + 172\lambda + 300)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{b^{10}\lambda^4(10\lambda^6 - 3\lambda^5 - 2\lambda + 3) - b^8\lambda^3(\lambda^6 - 54\lambda^5 - 12\lambda^4 - 20\lambda^3 + 26\lambda + 51)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{b^9\lambda^3(-15\lambda^7 - 72\lambda^6 - 9\lambda^5 - 4\lambda^4 + 6\lambda^2 + 10\lambda + 5) + b^{11}\lambda^4(\lambda^7 + 5\lambda^6 - 1)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{4b^4\lambda(10\lambda^7 + 147\lambda^6 + 287\lambda^5 - 12\lambda^4 - 18\lambda^3 - 22\lambda^2 - 122\lambda - 159)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{4b^3(12\lambda^7 - 9\lambda^6 + 257\lambda^5 - 36\lambda^4 - 62\lambda^3 + 24\lambda^2 + 18\lambda - 36)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{b^7\lambda^2(80\lambda^7 + 335\lambda^6 + 135\lambda^5 - 52\lambda^4 - 20\lambda^3 - 24\lambda^2 - 22\lambda - 92)}{4(b-2)^2(b+1)\lambda^4(b\lambda-3)^2(b\lambda+1)(b^2\lambda-6)^2(b^2\lambda-2)(b^2\lambda-1)} +$$

$$\frac{4b^5\lambda(2\lambda^8 + 33\lambda^7 + 105\lambda^6 + 151\lambda^5 - 30\lambda^4 - 41\lambda^3 + 8\lambda - 51)}{4(b-2)^2(b+1)\lambda^4(b\lambda - 3)^2(b\lambda + 1)(b^2\lambda - 6)^2(b^2\lambda - 2)(b^2\lambda - 1)} < \pi^{NL}$$

so again licensing will not occur and the potential innovator profit remains just that, potential profit that will not be realized.

APPENDIX B
BAYESIAN UPDATING

To illustrate how farmer update his prior belief on an uncertain input based on Bayes' rule, Let us consider Bayesian estimation of the mean of a univariate Gaussian with known variance.

Let D = $(x_1, \ldots, x_n)$ be the data. The likelihood is:

$$p(D|\mu,\sigma^2) = \prod_{i=1}^{n} p(x_i|\mu,\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

Mean and variance are:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2$$

the exponent term is:

$$\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}[(x_i-\bar{x})-(\mu-\bar{x})]^2$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + \sum_{i=1}^{n}(\mu-\bar{x})^2 - 2\sum_{i=1}^{n}(x_i-\bar{x})(\mu-\bar{x}) \quad = n\,s^2$$

$$+\, n(\mu-\bar{x})^2$$ [27]

Therefore

$$p(D|\mu,\sigma^2) = (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\left(-\frac{1}{2\sigma^2}n(\mu-\bar{x})^2 + ns^2\right)$$

$$\propto \frac{1}{\sigma^2}^{\frac{n}{2}} exp\left(-\frac{n}{2\sigma^2}(\mu-\bar{x})^2\right)\exp(-\frac{ns^2}{2\sigma^2})$$

Because σ^2 is constant

[27] $\sum(x_i-\bar{x})(\mu-\bar{x}) = (\mu-\bar{x})(\sum x_i - n\bar{x}) = (\mu-\bar{x})(n\bar{x}-n\bar{x}) = 0$

$$p(D|\mu, \sigma^2) \propto \; exp\left(-\frac{n}{2\sigma^2}(\mu - \bar{x})^2\right) \propto N(\bar{x}|\mu, \frac{\sigma^2}{n})$$

Assume the $p(\mu) \sim N(\mu|\mu_0, {\sigma^2}_0)$[28], which is the prior belief. The posterior belief after observing D = $(x_1, \ldots, x_n)$, $p(\mu|D)$ is:

$$p(\mu|D) \; \propto \; p(D|\mu, \sigma^2) \; p(\mu|\mu_0, {\sigma_0}^2)^{29}$$

$$\propto \; exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right) exp\left(-\frac{1}{2{\sigma_0}^2}(\mu - \mu_0)^2\right)$$

$$=\exp\left[-\frac{1}{2\sigma^2}(x_i - \mu)^2 - \frac{1}{2{\sigma_0}^2}(\mu - \mu_0)^2\right]$$

$$=\exp[-\frac{\mu^2}{2}\left(\frac{1}{{\sigma_0}^2} + \frac{n}{\sigma^2}\right) + \mu\left(\frac{\mu_0}{{\sigma_0}^2} + \frac{\sum_i x_i}{\sigma^2}\right) - (\frac{{\mu_0}^2}{2{\sigma_0}^2} + \frac{\sum_i {x_i}^2}{2\sigma^2})]$$

By definition, the posterior distribution is expressed as:

$$\exp\left[-\frac{1}{2\sigma_n^2}(\mu - \mu_n)^2\right] = \exp[-\frac{1}{2\sigma_n^2}(\mu^2 - 2\mu\mu_n + \mu_n^2)].$$

Matching the coefficients of μ^2, we find σ_n^2 to be:

$$\frac{1}{\sigma_n^2} = \frac{1}{{\sigma_0}^2} + \frac{n}{\sigma^2},$$

and that:

$$\sigma_n^2 = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{{\sigma_0}^2}}.$$

[28] ${\sigma_0}^2$ is the variance of the prior, while σ^2 is the variance of the observation noise.
[29] This is Bayes rule.

APPENDIX C.
EXPERIMENT INSTRUCTIONS

C1. Survey used for treatment groups.

Modeling Consumer Preferences for Innovative Products

You are being asked to take part in a study of consumer preferences regarding activity trackers. Please read this form carefully and ask any questions. You must be 18 years and older to participate.

What we will ask you to do: We will ask you to complete a set of questions regarding your workout habits, preferences for activity tracker attributes, and some demographic information such as age and gender. You will also be asked to talk about your choices during the experiment. The entire process should take about 30 minutes.

Risks: No additional risks other than those encountered in day-to-day life.

Compensation: You will receive $20 in cash before the experiment, upon our receipt of your signature.

Confidentiality: ALL answers but the tracker choices are confidential and anonymous.

Taking part is voluntary: Taking part in this study is completely voluntary. You may skip any questions that you do not want to answer. You are free to withdraw at any time.

If you have questions: The researchers conducting this study are Dr. Timothy Richards and Di Fang. Please ask any questions you have now. If you have questions later, you may contact Dr. Timothy Richards at trichards@asu.edu or Di Fang at dfang3@asu.edu or at 480-252-5931.

This research has been reviewed and approved by the Social Behavioral IRB. You may talk to them at (480) 965-6788 or by email at research.integrity@asu.edu.

Statement of Consent: I have read the above information, and have received answers to any questions I asked. I consent to take part in the study.

Your Signature ________________________________ Date ______________________

Your Name (printed) ____

Questionnaire Group No. _2____ Member No. _____

This is the first part of the survey. We would like to start with a few questions related to your workout frequency and sports related purchases. This is an anonymous survey and your name is not linked to the responses.

For the following questions, check or fill in the answers which best describe you.

1. How often do you workout, both inside and outside the gym?

Everyday	At least once every week	Once every other week	Once a month	Once a few months or less often

2. How often, on average, do you purchase sports related products (sports gear, apparel, etc.)?

At least once a week	Once a month	Once every three months	Once every six months	Once a year or less often

3. How much money do you spend on buying sports related goods every year?

Less than $100 ____
$100-$299 ____
$300-$499 ____
$500-$999 ____
$1,000 and more____

4. Do you currently own an activity tracker?

Yes____
No ____

Please stop here and wait for further information.

Questionnaire Group No. __2____ Member No. ______

Choice Survey

This is the second part of the survey. We would like to illustrate what an activity tracker is and its attributes. You will be asked to choose among different trackers. Your choices may not be anonymous but your name is not linked to the responses. We list some common activity tracker attributes below, and your choices will be made from among some combination of these attributes.

Brand	Fitbit Jawbone Nike Garmin
Design Clip on	  
Wristband	  
Watch	  
Function	*Recording basic calories--Cal* *Recording calories and sleep patterns—Cal+Sleep* *Recording calories and text/email messages—Cal+Msg* *Recording calories and GPS locations—Cal+GPS*
Price	$49.99

	$99.99 $129.99 $199.99

Sample Choice:

	A	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Nike	Fitbit	**None of These**
Design	Clip on	Watch	Wristband	Clip on	
Function	Cal+Sleep	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$129.99	$49.99	
Choice	□□□□□□□□				

IMPORTANT

- CHOOSE one of the options per table. Or you may choose None of These.
- Assume that the options in each table are the only ones available.
- Do not compare options across tables.

You might see a few options that are counter-intuitive (e.g., a lower price but a higher quality in your personal opinion). Be assured that this is not an error but part of the design of the survey. Simply choose the option that you prefer most.

Questionnaire Group No. _2____ Member No. _____

Choice Sets 1

Item# 1	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Nike	Fitbit	**None of These**
Design	Clip on	Clip on	Wristband	Clip on	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$49.99	$199.99	$129.99	$49.99	
Choice					

Item# 2	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Fitbit	Nike	Jawbone	**None of These**
Design	Clip on	Wristband	Wristband	Clip on	
Function	Cal+Sleep	Cal+Sleep	Cal+Sleep	Cal+Msg	
Price	$199.99	$199.99	$129.99	$49.99	
Choice					

Item# 3	**A**	**B**	**C**	**D**	**E**
Brand	Jawbone	Jawbone	Garmin	Fitbit	**None of These**
Design	Clip on	Watch	Watch	Watch	
Function	Cal	Cal	Cal	Cal+Msg	
Price	$49.99	$99.99	$129.99	$129.99	
Choice					

Item# 4	**A**	**B**	**C**	**D**	**E**
Brand	Fitbit	Nike	Nike	Fitbit	**None of These**
Design	Clip on	Wristband	Clip on	Clip on	
Function	Cal+GPS	Cal	Cal+Msg	Cal+Sleep	
Price	$99.99	$49.99	$99.99	$49.99	
Choice					

Item# 5	**A**	**B**	**C**	**D**	**E**
Brand	Fitbit	Fitbit	Garmin	Jawbone	**None of These**
Design	Clip on	Clip on	Watch	Wristband	
Function	Cal+Sleep	Cal+Msg	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$99.99	$49.99	
Choice					

Item# 6	**A**	**B**	**C**	**D**	**E**
Brand	Fitbit	Jawbone	Jawbone	Garmin	**None of These**
Design	Watch	Clip on	Wristband	Watch	

Function	Cal	Cal+Msg	Cal+GPS	Cal+Msg	
Price	$99.99	$99.99	$49.99	$49.99	
Choice					

Questionnaire Group No. __2____ Member No. ______

Item# 7	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Garmin	Jawbone	Jawbone	**None of These**
Design	Wristband	Clip on	Clip on	Clip on	
Function	Cal+Sleep	Cal+Msg	Cal+Sleep	Cal+Sleep	
Price	$199.99	$129.99	$129.99	$199.99	
Choice					

Item# 8	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Jawbone	Garmin	**None of These**
Design	Wristband	Wristband	Watch	Watch	
Function	Cal	Cal+Msg	Cal+Sleep	Cal+GPS	
Price	$199.99	$199.99	$49.99	$199.99	
Choice					

Item# 9	**A**	**B**	**C**	**D**	**E**
Brand	Nike	Jawbone	Fitbit	Garmin	**None of These**
Design	Wristband	Clip on	Wristband	Wristband	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+GPS	
Price	$199.99	$129.99	$99.99	$129.99	
Choice					

Item# 10	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Garmin	Nike	Nike	**None of These**
Design	Clip on	Watch	Clip on	Watch	
Function	Cal+Sleep	Cal	Cal+Sleep	Cal	
Price	$129.99	$49.99	$49.99	$129.99	
Choice					

Item# 11	**A**	**B**	**C**	**D**	**E**
Brand	Jawbone	Fitbit	Nike	Garmin	**None of These**
Design	Wristband	Wristband	Watch	Clip on	
Function	Cal	Cal+GPS	Cal+GPS	Cal	
Price	$129.99	$129.99	$99.99	$99.99	
Choice					

Item# 12	**A**	**B**	**C**	**D**	**E**
Brand	Nike	Garmin	Jawbone	Nike	

Design	Clip on	Wristband	Watch	Wristband	None of These
Function	Cal	Cal	Cal+GPS	Cal+GPS	
Price	$199.99	$99.99	$199.99	$49.99	
Choice					

If you've reached here please stop and wait for further information.

Questionnaire Group No. __2___ Member No. _____

5. Please indicate below how well you know the other participants.

1 = have never met
2 = met once
3 = somewhat acquainted
4 = well acquainted
5 = know him/her well

For example if you know participant #1 well, then put number 5 next to #1 (as #1_5_); if you have never met participant #1 then put 1 next to #1(as #1_1_).

#1___

#2___

#3___

#4___

#5___

#6___

#7___

#8___

#9___

#10___

#11___

#12___

#13___

#14___

#15___

#16___

#17___

#18___

#19___

#20___

Let's talk about your choices!

Questionnaire Group No. ____2__ Member No. ______

Choice experiments 2

Item# 1	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Garmin	Jawbone	Jawbone	**None of These**
Design	Wristband	Clip on	Clip on	Clip on	
Function	Cal+Sleep	Cal+Msg	Cal+Sleep	Cal+Sleep	
Price	$199.99	$129.99	$129.99	$199.99	
Choice					

Item# 2	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Jawbone	Garmin	**None of These**
Design	Wristband	Wristband	Watch	Watch	
Function	Cal	Cal+Msg	Cal+Sleep	Cal+GPS	
Price	$199.99	$199.99	$49.99	$199.99	
Choice					

Item# 3	**A**	**B**	**C**	**D**	**E**
Brand	Nike	Jawbone	Fitbit	Garmin	**None of These**
Design	Wristband	Clip on	Wristband	Wristband	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+GPS	
Price	$199.99	$129.99	$99.99	$129.99	
Choice					

Item# 4	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Garmin	Nike	Nike	**None of These**
Design	Clip on	Watch	Clip on	Watch	
Function	Cal+Sleep	Cal	Cal+Sleep	Cal	
Price	$129.99	$49.99	$49.99	$129.99	
Choice					

Item# 5	**A**	**B**	**C**	**D**	**E**

Brand	Jawbone	Fitbit	Nike	Garmin	**None of These**
Design	Wristband	Wristband	Watch	Clip on	
Function	Cal	Cal+GPS	Cal+GPS	Cal	
Price	$129.99	$129.99	$99.99	$99.99	
Choice					

Item# 6	A	B	C	D	E
Brand	Nike	Garmin	Jawbone	Nike	**None of These**
Design	Clip on	Wristband	Watch	Wristband	
Function	Cal	Cal	Cal+GPS	Cal+GPS	
Price	$199.99	$99.99	$199.99	$49.99	
Choice					

Questionnaire Group No. ___2_ Member No. _____

Item# 7	A	B	C	D	E
Brand	Garmin	Jawbone	Nike	Fitbit	**None of These**
Design	Clip on	Clip on	Wristband	Clip on	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$49.99	$199.99	$129.99	$49.99	
Choice					

Item# 8	A	B	C	D	E
Brand	Garmin	Fitbit	Nike	Jawbone	**None of These**
Design	Clip on	Wristband	Wristband	Clip on	
Function	Cal+Sleep	Cal+Sleep	Cal+Sleep	Cal+Msg	
Price	$199.99	$199.99	$129.99	$49.99	
Choice					

Item# 9	A	B	C	D	E
Brand	Jawbone	Jawbone	Garmin	Fitbit	**None of These**
Design	Clip on	Watch	Watch	Watch	
Function	Cal	Cal	Cal	Cal+Msg	
Price	$49.99	$99.99	$129.99	$129.99	
Choice					

Item# 10	A	B	C	D	E
Brand	Fitbit	Nike	Nike	Fitbit	**Not Interested**
Design	Clip on	Wristband	Clip on	Clip on	
Function	Cal+GPS	Cal	Cal+Msg	Cal+Sleep	
Price	$99.99	$49.99	$99.99	$49.99	
Choice					

Item# 11	A	B	C	D	E
Brand	Fitbit	Fitbit	Garmin	Jawbone	**None of These**
Design	Clip on	Clip on	Watch	Wristband	
Function	Cal+Sleep	Cal+Msg	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$99.99	$49.99	
Choice					

Item# 12	A	B	C	D	E
Brand	Fitbit	Jawbone	Jawbone	Garmin	**None of These**
Design	Watch	Clip on	Wristband	Watch	
Function	Cal	Cal+Msg	Cal+GPS	Cal+Msg	
Price	$99.99	$99.99	$49.99	$49.99	
Choice					

Questionnaire Group No. _2____ Member No. _____

6. Please indicate below how you rate the credibility of the opinions offered by other people in your group.
1 = not credible
2 = a small amount of credibility
3 = indifferent
4 = somewhat credible
5 = very credible
For example you value participant #1's opinion very much then put 5 next to #1 (as #1_5_); if you do not have respect for #1's opinion then put 1 next to #1 (as #1_1_).

#1___

#2___

#3___

#4___

#5___

#6___

#7___

#8___

#9___

#10___

#11___

#12___

#13___

#14___

#15___

#16___

#17___

#18___

#19___

#20___

Questionnaire Group No. __2____ Member No. ______

This is the final part of the survey. I would like to ask you a few questions about yourself. As mentioned in the beginning, this is an anonymous survey and your name is not linked to the responses. Please remember that you can skip as many questions as you want or withdraw at anytime.

7. How old are you?

_________years

8. Please indicate your gender

Male ( ) Female ( )

9. Please indicate which ethnic group you belong to?

White _____ Hispanic ____ Native American ____
African American _____ Asian/Pacific Islander _____ Other ______________

10. What year are you in your program?

First year _______
Second year _______
Third year _______
Fourth year and more_______
Graduate student_______

Other_____

11. Please indicate your weight and height – make your best guess.

Weight indicate lb _________ Height indicate inches ______

12. Please indicate your approximate annual household income before taxes:

Less than $10,000	_____	$60,000 to $69,999	_____
$10,000 to $19,999	_____	$70,000 to $79,999	_____
$20,000 to $29,999	_____	$80,000 to $89,999	_____
$30,000 to $39,999	_____	$90,000 to $99,999	_____
$40,000 to $49,999	_____	$100, 000 to $149,999	_____
$50,000 to $59,999	_____	$150,000 or more	_____

Thank you very much for your participation!

C2. Survey for control groups.

Modeling Consumer Preferences for Innovative Products

You are being asked to take part in a study of consumer preferences regarding activity trackers. Please read this form carefully and ask any questions. You must be 18 years and older to participate.

What we will ask you to do: We will ask you to complete a set of questions regarding your workout habits, preferences for activity tracker attributes, and some demographic information such as age and gender. You will read a short article from New York Times. The entire process should take about 30 minutes.

Risks: No additional risks other than those encountered in day-to-day life.

Compensation: You will receive $20 in cash before the experiment, upon our receipt of your signature.

Confidentiality: ALL answers are confidential and anonymous.

Taking part is voluntary: Taking part in this study is completely voluntary. You may skip any questions that you do not want to answer. You are free to withdraw at any time.

If you have questions: The researchers conducting this study are Dr. Timothy Richards and Di Fang. Please ask any questions you have now. If you have questions later, you may contact Dr. Timothy Richards at trichards@asu.edu or Di Fang at dfang3@asu.edu or at 480-252-5931.

This research has been reviewed and approved by the Social Behavioral IRB. You may talk to them at (480) 965-6788 or by email at research.integrity@asu.edu.

Statement of Consent: I have read the above information, and have received answers to any questions I asked. I consent to take part in the study.

Your Signature ____________________________________ Date ________________________

Your Name (printed) ____

Questionnaire Group No. __5___ Member No. ______

This is the first part of the survey. We would like to start with a few questions related to your workout frequency and sports related purchases. This is an anonymous survey and your name is not linked to the responses.

For the following questions, check or fill in the answers which best describe you.

1. How often do you workout, both inside and outside the gym?

Everyday	At least once every week	Once every other week	Once a month	Once a few months or less often

2. How often, on average, do you purchase sports related products (sports gear, apparel, etc.)?

At least once a week	Once a month	Once every three months	Once every six months	Once a year or less often

3. How much money do you spend on buying sports related goods every year?

Less than $100 _____
$100-$299 _____
$300-$499 _____
$500-$999 _____
$1,000 and more_____

4. Do you currently own an activity tracker?

Yes_____
No _____

Questionnaire Group No. ___5___   Member No. _____

Choice Survey

This is the second part of the survey. We would like to illustrate what an activity tracker is and its attributes. You will be asked to choose among different trackers. Your choices may not be anonymous but your name is not linked to the responses. We list some common activity tracker attributes below, and your choices will be made from among some combination of these attributes.

Brand	Fitbit Jawbone Nike Garmin
Design Clip on	  
Wristband	  
Watch	  
Function	*Recording basic calories--Cal* *Recording calories and sleep patterns—Cal+Sleep* *Recording calories and text/email messages—Cal+Msg* *Recording calories and GPS locations—Cal+GPS*
Price	$49.99 $99.99

	$129.99 $199.99

Sample Choice:

	A	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Nike	Fitbit	**None of These**
Design	Clip on	Watch	Wristband	Clip on	
Function	Cal+Sleep	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$129.99	$49.99	
Choice	□□□□□□□	□□□□□□□✓			

IMPORTANT

- CHOOSE one of the options per table. Or you may choose None of These.
- Assume that the options in each table are the only ones available.
- Do not compare options across tables.

You might see a few options that are counter-intuitive (e.g., a lower price but a higher quality in your personal opinion). Be assured that this is not an error but part of the design of the survey. Simply choose the option that you prefer most.

Item# 1	A	B	C	D	E
Brand	Garmin	Jawbone	Nike	Fitbit	None of These
Design	Clip on	Clip on	Wristband	Clip on	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$49.99	$199.99	$129.99	$49.99	
Choice					

Item# 2	A	B	C	D	E
Brand	Garmin	Fitbit	Nike	Jawbone	None of These
Design	Clip on	Wristband	Wristband	Clip on	
Function	Cal+Sleep	Cal+Sleep	Cal+Sleep	Cal+Msg	
Price	$199.99	$199.99	$129.99	$49.99	
Choice					

Item# 3	A	B	C	D	E
Brand	Jawbone	Jawbone	Garmin	Fitbit	None of These
Design	Clip on	Watch	Watch	Watch	
Function	Cal	Cal	Cal	Cal+Msg	
Price	$49.99	$99.99	$129.99	$129.99	
Choice					

Item# 4	A	B	C	D	E
Brand	Fitbit	Nike	Nike	Fitbit	None of These
Design	Clip on	Wristband	Clip on	Clip on	
Function	Cal+GPS	Cal	Cal+Msg	Cal+Sleep	
Price	$99.99	$49.99	$99.99	$49.99	
Choice					

Item# 5	A	B	C	D	E
Brand	Fitbit	Fitbit	Garmin	Jawbone	None of These
Design	Clip on	Clip on	Watch	Wristband	
Function	Cal+Sleep	Cal+Msg	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$99.99	$49.99	
Choice					

Item# 6	A	B	C	D	E
Brand	Fitbit	Jawbone	Jawbone	Garmin	None of These
Design	Watch	Clip on	Wristband	Watch	
Function	Cal	Cal+Msg	Cal+GPS	Cal+Msg	
Price	$99.99	$99.99	$49.99	$49.99	
Choice					

Questionnaire Group No. __5____ Member No. ______

Item# 7	A	B	C	D	E
Brand	Garmin	Garmin	Jawbone	Jawbone	None of These
Design	Wristband	Clip on	Clip on	Clip on	
Function	Cal+Sleep	Cal+Msg	Cal+Sleep	Cal+Sleep	
Price	$199.99	$129.99	$129.99	$199.99	
Choice					

Item# 8	A	B	C	D	E
Brand	Garmin	Jawbone	Jawbone	Garmin	None of These
Design	Wristband	Wristband	Watch	Watch	
Function	Cal	Cal+Msg	Cal+Sleep	Cal+GPS	
Price	$199.99	$199.99	$49.99	$199.99	
Choice					

Item# 9	A	B	C	D	E
Brand	Nike	Jawbone	Fitbit	Garmin	None of These
Design	Wristband	Clip on	Wristband	Wristband	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+GPS	
Price	$199.99	$129.99	$99.99	$129.99	
Choice					

Item# 10	A	B	C	D	E
Brand	Garmin	Garmin	Nike	Nike	None of These
Design	Clip on	Watch	Clip on	Watch	
Function	Cal+Sleep	Cal	Cal+Sleep	Cal	
Price	$129.99	$49.99	$49.99	$129.99	
Choice					

Item# 11	A	B	C	D	E
Brand	Jawbone	Fitbit	Nike	Garmin	None of These
Design	Wristband	Wristband	Watch	Clip on	
Function	Cal	Cal+GPS	Cal+GPS	Cal	
Price	$129.99	$129.99	$99.99	$99.99	
Choice					

Item# 12	A	B	C	D	E
Brand	Nike	Garmin	Jawbone	Nike	None of These
Design	Clip on	Wristband	Watch	Wristband	
Function	Cal	Cal	Cal+GPS	Cal+GPS	
Price	$199.99	$99.99	$199.99	$49.99	
Choice					

Questionnaire Group No. __5___ Member No. _____

5. Please indicate below how well you know the other participants.
1 = have never met
2 = met once
3 = somewhat acquainted
4 = well acquainted
5 = know him/her well
For example if you know participant #1 well, then put number 5 next to #1 (as #1_5_); if you have never met participant #1 then put 1 next to #1(as #1_1_).

#1___

#2___

#3___

#4___

#5___

#6___

#7___

#8___

#9___

#10___

#11___

#12___

#13___

#14___

#15___

#16___

#17___

#18___

#19___

#20___

Please read the following article from New York Times

China Acts to Stem Slide in Home Prices

MARCH 30, 2015

SHANGHAI — China on Monday courted home buyers with a bigger tax break as it cut down-payment requirements for the second time in six months, stepping up a fight against sliding house prices that is imperiling the Chinese economy.

The People's Bank of China, the central bank, said on its website that commercial banks could now lower their minimum down-payment requirement for buyers of second homes, and with outstanding mortgages, to 40 percent from 60 percent.

The Ministry of Finance, in a separate statement, said that individuals selling houses were exempt from business taxes if they had owned the house for more than two years. Analysts said sellers were previously exempted from taxes only if they owned the houses for at least five years.

The policy sweeteners, which were more generous than what the market had expected, confirmed rumors swirling in China on Monday that the authorities were increasing support for the flagging real estate sector. Real estate share indexes rallied sharply in Shanghai on rumors of the change. The Shanghai composite's property index closed up more than 7 percent, its best day since 2009, while the broader index closed 2.6 percent higher.

That China is now trying to lift its property market is an about-face in policy. As recently as early 2014, the authorities were waging a four-year campaign to tame an exuberant market, which pushed home prices to records.

Some analysts doubted that the measures announced Monday would lead to a turnaround.

"The new measures are definitely helpful, but the impact won't be significant," said Ada Wong, a vice president at the China Aoyuan Property Group, a developer based in Guangzhou. "This is because most of the speculative buyers have been eliminated. The market has become more rational, so sales won't increase a lot."

Zhu Haibin, an economist at JPMorgan Chase, said, "We expect the housing market correction will continue, but at a relatively modest pace through the course of this year. In our view, real estate investment growth will likely further decelerate from 10.5 percent in 2014 to about 6 percent in 2015, which will continue to drag on economic growth."

The housing market has increasingly weighed on the economy. Prices fell at a record pace in February, denting activity in industries like cement, steel and glass making.

Real estate accounts for about 15 percent of the Chinese economic activity and economists have warned that persistent weakness in housing could endanger

Beijing’s target of 7 percent growth this year.

Investors had speculated that China would cut down payments for buyers of second homes with outstanding home loans to 50 percent. Yet while the move was more generous than many had expected, it was not clear that banks would pass on the discounts to buyers.

Even before the latest relaxation, some banks were demanding down payments of as much as 70 percent, higher than a government-set minimum of 60 percent.

In its statement, the central bank urged financial institutions to “support home purchases” with a combination of commercial lending and money from the national provident fund. The provident fund is a government-managed program in which all of China’s employers and employees contribute to a pool of money from which employees then borrow to buy homes.

The central bank said the down payment for second-home loans by borrowers from the housing provident fund was now set at 30 percent, provided borrowers had no outstanding mortgages. First-time buyers using the provident fund need make a down payment of only 20 percent, the bank said, with all changes taking immediate effect.

Please continue with the survey, thank you so much for your patience!

Questionnaire Group No. ____5__ Member No. _____

Choice experiments 2

Item# 1	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Garmin	Jawbone	Jawbone	**None of These**
Design	Wristband	Clip on	Clip on	Clip on	
Function	Cal+Sleep	Cal+Msg	Cal+Sleep	Cal+Sleep	
Price	$199.99	$129.99	$129.99	$199.99	
Choice					

Item# 2	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Jawbone	Garmin	**None of These**
Design	Wristband	Wristband	Watch	Watch	
Function	Cal	Cal+Msg	Cal+Sleep	Cal+GPS	
Price	$199.99	$199.99	$49.99	$199.99	
Choice					

Item# 3	**A**	**B**	**C**	**D**	**E**
Brand	Nike	Jawbone	Fitbit	Garmin	**None of These**
Design	Wristband	Clip on	Wristband	Wristband	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+GPS	
Price	$199.99	$129.99	$99.99	$129.99	
Choice					

Item# 4	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Garmin	Nike	Nike	**None of These**
Design	Clip on	Watch	Clip on	Watch	
Function	Cal+Sleep	Cal	Cal+Sleep	Cal	
Price	$129.99	$49.99	$49.99	$129.99	
Choice					

Item# 5	**A**	**B**	**C**	**D**	**E**
Brand	Jawbone	Fitbit	Nike	Garmin	**None of These**
Design	Wristband	Wristband	Watch	Clip on	
Function	Cal	Cal+GPS	Cal+GPS	Cal	
Price	$129.99	$129.99	$99.99	$99.99	
Choice					

Item# 6	**A**	**B**	**C**	**D**	**E**
Brand	Nike	Garmin	Jawbone	Nike	

Design	Clip on	Wristband	Watch	Wristband	**None of These**
Function	Cal	Cal	Cal+GPS	Cal+GPS	
Price	$199.99	$99.99	$199.99	$49.99	
Choice					

Questionnaire Group No. ____5_ Member No. _____

Item# 7	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Jawbone	Nike	Fitbit	**None of These**
Design	Clip on	Clip on	Wristband	Clip on	
Function	Cal+GPS	Cal+GPS	Cal+Msg	Cal+Msg	
Price	$49.99	$199.99	$129.99	$49.99	
Choice					

Item# 8	**A**	**B**	**C**	**D**	**E**
Brand	Garmin	Fitbit	Nike	Jawbone	**None of These**
Design	Clip on	Wristband	Wristband	Clip on	
Function	Cal+Sleep	Cal+Sleep	Cal+Sleep	Cal+Msg	
Price	$199.99	$199.99	$129.99	$49.99	
Choice					

Item# 9	**A**	**B**	**C**	**D**	**E**
Brand	Jawbone	Jawbone	Garmin	Fitbit	**None of These**
Design	Clip on	Watch	Watch	Watch	
Function	Cal	Cal	Cal	Cal+Msg	
Price	$49.99	$99.99	$129.99	$129.99	
Choice					

Item# 10	**A**	**B**	**C**	**D**	**E**
Brand	Fitbit	Nike	Nike	Fitbit	**Not Interested**
Design	Clip on	Wristband	Clip on	Clip on	
Function	Cal+GPS	Cal	Cal+Msg	Cal+Sleep	
Price	$99.99	$49.99	$99.99	$49.99	
Choice					

Item# 11	**A**	**B**	**C**	**D**	**E**
Brand	Fitbit	Fitbit	Garmin	Jawbone	**None of These**
Design	Clip on	Clip on	Watch	Wristband	
Function	Cal+Sleep	Cal+Msg	Cal+Msg	Cal+Msg	
Price	$99.99	$199.99	$99.99	$49.99	
Choice					

Item# 12	**A**	**B**	**C**	**D**	**E**

Brand	Fitbit	Jawbone	Jawbone	Garmin	**None of These**
Design	Watch	Clip on	Wristband	Watch	
Function	Cal	Cal+Msg	Cal+GPS	Cal+Msg	
Price	$99.99	$99.99	$49.99	$49.99	
Choice					

Questionnaire Group No. __5___ Member No. ______

This is the final part of the survey. I would like to ask you a few questions about yourself. As mentioned in the beginning, this is an anonymous survey and your name is not linked to the responses. Please remember that you can skip as many questions as you want or withdraw at anytime.

7. How old are you?

______years

8. Please indicate your gender

Male ( ) Female ( )

9. Please indicate which ethnic group you belong to?

White _____ Hispanic ____ Native American ____
African American _____ Asian/Pacific Islander _____ Other _______________

10. What year are you in your program?

First year _______
Second year _______
Third year _______
Fourth year and more_______
Graduate student_______
Other______

11. Please indicate your weight and height – make your best guess.
Weight indicate lb ____________ Height indicate inches _______

12. Please indicate your approximate annual household income before taxes:

Less than $10,000 ______ $60,000 to $69,999 ______
$10,000 to $19,999 ______ $70,000 to $79,999

____

$20,000 to $29,999	_______	$80,000 to $89,999	_______
$30,000 to $39,999	_______	$90,000 to $99,999	_______
$40,000 to $49,999	_______	$100, 000 to $149,999	_______
$50,000 to $59,999	_______	$150,000 or more	_______

Thank you very much for your participation!

C2. Recruiting Scripts using Qualtrics

Default Question Block

Dear Participant:

My name is Di Fang. I am a PhD student in the Morrison School of Agribusiness, W.P. Carey School of Business.

I am conducting a study about **activity trackers** (Fitbit, Nike, Jawbone, Garmin) choices ***from April 1st-April 30***. This will take approximately ***30 minutes*** and you will be given ***$20*** in cash for your participation.

During the study, you will be asked to finish a survey about your demographic information(age, gender, etc). You will also be asked to choose among different activity trackers and maybe (depend on which group you are assigned) discuss your choices with others. Your participation is voluntary and you can skip any question you do not wish to answer. You need to be 18 years or older.

If you like to participate please fill out the next couple questions and pick out a time slot. Feel free to text me at (480) 252-5931 or email dfang3@asu.edu if you have questions.

Thanks for your time!
Di Fang

Please make sure to fill out your **department** and **major** and leave an **email** address so I can contact you about time of the experiment.

Name	
Department/College	
Email	
Major	
How did you learn about this study(please be specific)?	

On which campus can you attend this study?

- Tempe (main) campus
- Polytechnic (east) campus
- Both

Are you 18 and older?

- Yes
- No

What year are you?

First year

Second year

Third year

Fourth year and more

Graduate students

For scheduling purpose, please list two preferred day and time, for example Mon 8am-10am. Remember this study takes about 30 minutes.

Day and Time 1	
Day and Time 2	
If you want to be assigned to the same group with your friends please indicate their names.	

Thank you for your time and patience. I will email/call you back if you are scheduled successfully!

Please forward this email to your friends so they can come too!

C3. IRB Approval Letter

ASU Knowledge Enterprise Development

EXEMPTION GRANTED

Timothy Richards
Agribusiness, Morrison School of
480/727-1488
trichards@asu.edu

Dear Timothy Richards:

On 3/4/2015 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Modeling Consumer Preferences for Innovative Product in a Friendship Network
Investigator:	Timothy Richards
IRB ID:	STUDY00002226
Funding:	Name: Agribusiness, Morrison School of; WPC
Grant Title:	
Grant ID:	
Documents Reviewed:	• IRB Protocol, Category: IRB Protocol; • Consent Form_Revised March 3.pdf, Category: Consent Form; • Flyer, Category: Recruitment materials/advertisements /verbal scripts/phone scripts; • Grant Scope, Category: Other (to reflect anything not captured above); • Complete Profile and Sample Card_Revised March 3.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Recruiting Script.pdf, Category: Recruitment Materials; • Friendship Rating Document, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Demographic Questionnaire.pdf, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions);