DISSERTATION

REGIONAL DIMENSIONS OF AGRITOURISM: EXPLORING SPATIAL AND TRAVELER HETEROGENEITY

Submitted by

Anders Van Sandt

Department of Agricultural and Resource Economics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2018

Doctoral Committee:

Advisor: Dawn Thilmany McFadden

Marco Costanigro
Becca Jablonski
Stephan Weiler

ABSTRACT

REGIONAL DIMENSIONS OF AGRITOURISM: EXPLORING SPATIAL AND TRAVELER HETEROGENEITY

This dissertation employs a three-pronged approach to explore how locational, firm, and traveler heterogeneity lead to different opportunities or barriers in the budding U.S. agritourism industry. While each chapter considers a different aspect of the agritourism industry using unique empirical methods of analyses and sources of data, each chapter utilizes spatial economic methods to analyze different aspects of agritourism in the U.S. The first chapter applies firm level data and a two-stage model to test the importance of three trade theories in identifying comparative advantages in agritourism. Findings from this first essay imply that while some comparative advantages may be due to the productive and technological efficiencies of an agricultural operation, locational characteristics such as natural endowments (including natural amenities, farm production type, and proximities to other tourism attractions and population centers) are the primary source of attraction for agritourists. This essay also finds strong evidence of economies of agglomeration within the agritourism industry, signaling the need for future research to explore the potential within- and between-industry benefits from developing agritourism clusters.

The second essay estimates the consumer surplus derived from agritourism in the Western U.S. using a flexible travel cost model and survey data. In addition to providing par-worth consumer surplus measures across multiple regions, agritourism activities, and travelers, the method includes a detailed examination of how sunk costs of primary travel destinations may

influence consumer welfare estimates for other site visits on trips. Findings show that this mismeasurement inflicts a bias, called the multi-destination bias, that differs depending on the relative price and rurality of the recreational activity. Finally, the chapter includes a discussion on how relative elasticities may be used by agritourism operators and rural economic development practitioners to leverage their locational and site specific comparative advantages in agritourism.

The third and final essay analyzes primary data from a choice experiment in a latent class logit framework to investigate how consumers' home surroundings influence their willingness to pay for various agritourism qualities when choosing among destinations. After identifying each of the market segments and how they differ in regards to their agritourism preferences, the membership covariates are used to create willingness to pay maps using kriging, a geostatistical interpolation method. Maps and willingness to pay estimates from this analysis may be used by agritourism operators and tourism development practitioners to target marketing efforts in regions with significantly higher willingness to pay values.

By understanding how agritourism demand and supply factors differ, farmers and ranchers will be able to identify and leverage their natural, firm, and community strengths to develop successful agritourism enterprises. While agritourism is a relatively well-established industry in Europe, research opportunities to inform the US sector still exist. From the diversity of American agricultural producers, to exploring the potential spillover benefits to communities, and the demand shifts that may arise alongside the U.S. public's growing interest in food makes, there are several motivations for further research on this. Directions for future research are outlined at the end of each essay, as well as at the end of the conclusion.

ACKNOWLEDGMENTS

When people think of getting a Ph.D. they usually consider the skills, knowledge, and opportunities gained from earning a higher degree, but less often do people consider the impact this journey leaves on their personal growth and identity. Not only have my advisor, committee, and family and friends assisted me in receiving my degree, but their unwavering support has led me to emerge from the other side as a more confident, compassionate, and independent person. While my five years in graduate school were filled with plenty of mental and emotional challenges, their guidance and friendships have filled me with courage, light, and love.

By some stroke of luck, I was fortunate enough to be paired with my advisor, Dawn, since my first day at Colorado State University. There is no one else I would rather have fighting in my corner than Dawn. Her tenacity, brilliance, humility, and generosity have made her not only the greatest mentor a pupil could ask for, but also a cherished friend. When I reflect on the factors that led to my success in graduate school, her guidance, humor, and positive energy are at the top of the list, and I am confident I would not have experienced the same level of success or joy working with anyone else.

Given my stubbornness and aptitude for adventure, I have not been the easiest child to raise, and would like to thank my parents for their love, patience, and wisdom (regardless of whether or not I listened to their advice). Their frequent phone calls, cards, and visits out to Colorado have filled me with the love and courage to press on in my darkest times and have enriched my everyday life. Despite being 1,200 miles from home, I frequently find myself smiling and thinking about how lucky I am to have such caring, devoted, and at times goofy parents.

Finally, I would like to thank my friends and colleagues. Although not explicitly on my committee, Sarah Low, has been an extremely influential mentor. Sarah took me under her wing and helped me develop professionally through an internship in Washington, D.C., networking at conferences, and in exploring regional economics. My roommates and fiends, Chippy and Chompo, were two of the best distractions from the everyday slog, and kept me sane via late night board games, growlers of craft beer, and ultimate frisbee. Last and not least, I would like to thank Ash Heim for her patience and support during my brain injury, her companionship on our countless adventures, and her contagious determination. These people along with countless others in the Fort Collins community have shaped my time in graduate school and have inspired me to always keep improving who I am, both professionally and personally.

TABLE OF CONTENTS

ABSTRACT ii
ACKNOWLEDGMENTS iv
CHAPTER ONE: INTRODUCTION 1
CHAPTER TWO: PLACE-BASED FACTORS AND THE SUCCESS OF FARM-LEVEL ENTREPRENEURSHIP 12
- Introduction 12
- Literature Review 16
  - Identifying Sources of Comparative Advantage 20
- Methods 23
  - Heckscher-Ohlin 29
  - Ricardian 31
  - Agglomeration 32
- Empirical Results and Discussion 33
  - Propensity to Adopt Agritourism 37
  - Pull-factors of Agritourism Establishments 41
- Conclusions and Future Research 45

CHAPTER THREE: EXAMINING MULTI-DESTINATION BIAS IN HETEROGENEOUS AGRITOURISM-BASED TRAVEL COST MODELS 49
- Introduction 49
- Literature Review 53
- Methods 55
  - Accounting for Industry Heterogeneity 61
  - Empirical Model 65
  - Survey Data 66
- Results 73
- Conclusions 84

CHAPTER FOUR: THE IMPACT OF CONSUMER AND LOCATION HETEROGENEITY IN DETERMINING TRAVELER BEHAVIOR 88
- Introduction 88
- Methods 94
  - Survey Data 94
  - Theory and Empirical Model 97
  - Spatial Interpolation 102
- Empirical Results 104
  - Spatial Interpolation 106
- Discussion and Conclusion 110

CHAPTER FIVE: CONCLUSIONS 116
REFERENCES 124
APPENDIX 134
- Technical Appendix 137

CHAPTER ONE: INTRODUCTION

American agriculture and the diverse rural communities that principally rely on ag-based industries have undergone a major structural shift in recent decades. However, given the geographic scope of the U.S. and varied regional sub-cultures, not all communities have responded in the same fashion or even experienced the same set of market pressures. In general, as agriculture becomes relatively more capital intensive, the sector has made significant investments in labor-saving innovations to meet the challenges of globalized markets (based on green revolution advances and mechanization). As a result, many rural communities have experienced lagging employment growth and subsequent out-migration of young and educated residents (USDA, 2018). Concerns about this transition in agricultural based rural economies, and the resulting negative impacts on job opportunities and rural economies, have motivated some farms and ranches to diversify their agricultural businesses using a variety of alternative marketing strategies that leverage their farm and locational attributes. Agritourism is one commonly adopted strategy for farms due to the additional stream of revenue it provides, allowing employment opportunities for a family member, as it may serve to educate the public about agriculture and build a base of support in the surrounding community.

Between 2007 and 2012, the number of agricultural operators reporting agritourism revenues grew on net 42% in the U.S., in large part due to increasing commodity market pressures and the versatility of agritourism to act as a diversification strategy (USDA, 2014a; Veeck, G., Che, & A. Veeck, 2006). The relative share of farms and ranches with agritourism also grew 48% in the same period, but the share of farms and ranches with agritourism can vary drastically from state to state and even county to county. In addition to agritourism, emerging market challenges and a rising public interest in food have encouraged a set of small-medium sized agricultural producers to adopt other, complementary diversification strategies, such as niche market production and labeling (e.g. local, organic, natural, etc.), direct to consumer sales (e.g. farmstands and farmers markets), and direct to retailer sales (e.g. farm to table). Among alternative marketing strategies, agritourism stands out as a particularly interesting diversification strategy to explore due to the fact that it can be integrated into a large variety of agricultural production types, can be customized to attract travelers across most regions of the U.S., and moreover, its potential to act as a broader rural community development tool.

Considering the growth in the number of farms and ranches adopting agritourism across a diverse set of US agricultural producers, the factors driving adoption and success of such enterprises remains a relatively underdefined and understudied topic. This dissertation presents three essays that seek to fill gaps in the literature by answering the following research questions:

1) How do operational characteristics and place-based factors relate to a farm's propensity to adopt and level of economic activity tied to an agritourism enterprise?
2) How does consumers' travel behavior inform the sector about the value they place on agritourism experiences across multiple regions and activities?
3) How is heterogeneity in consumer preferences influenced by the consumers' location of origin, and how does willingness to pay for specific agritourism qualities differ depending on the location of the agritourism activity?

While each chapter is intended to make specific individual empirical or methodological contributions to the literature, they collectively present and develop ideas that may assist agricultural producers, policy makers, and rural development practitioners in developing a set of targeted business development strategies and best practices that take into account the nuances of this industry across space. However, this variability in the agritourism sector also presents challenges to studying the industry, requiring unique and flexible research methods to draw usable inferences.

Agritourism is most commonly defined in the literature as any revenue generating recreational or educational activity on a working farm or ranch that consumers pay to participate in. Examples of agritourism activities includes direct to consumer sales at the farm site, such as a farm stand, but also, farm-based entertainment and events, such as weddings and farm dinners, as well as outdoor recreation such as hunting or fishing or educational activities like petting zoos or food preparation classes. Clearly, this broad definition of agritourism indicates that it is an industry that differs both across activities and space with no unified product space, and this creates challenges for researchers interested in assessing its usefulness as a diversification strategy and possible rural development tool. However, studying the nuances between travelers, sites, and locational factors will not only assist in informing relevant industry stakeholders of best practices and market opportunities, but may also provide research approaches and generalized inferences that can be extended to the literature in support of other place-based industries.

Current agritourism patterns across the US motivated the questions explored in this study. Figures 1.1 and 1.2 show the number of agritourism farms and the percent of farms with agritourism revenue per county, respectively. The Pacific coast, Rocky Mountains, Texas, and the Northeast have higher concentrations of agritourism establishments while the Midwest and South have relatively few agritourism establishments. Even when controlling for the total number of farms and ranches, the Rocky Mountains, Texas, and the Northeast still have a higher degree of agritourism activity relative to other regions. Somehow, despite the variety of crops and cultures they encompass, certain counties in the U.S. appear to be more conducive to spawning agritourism activities than others, suggesting a potential clustering effect in this sector. So, it is worth exploring how the determinants of establishment and revenues in this industry may vary by traveler, site, and locational factors.

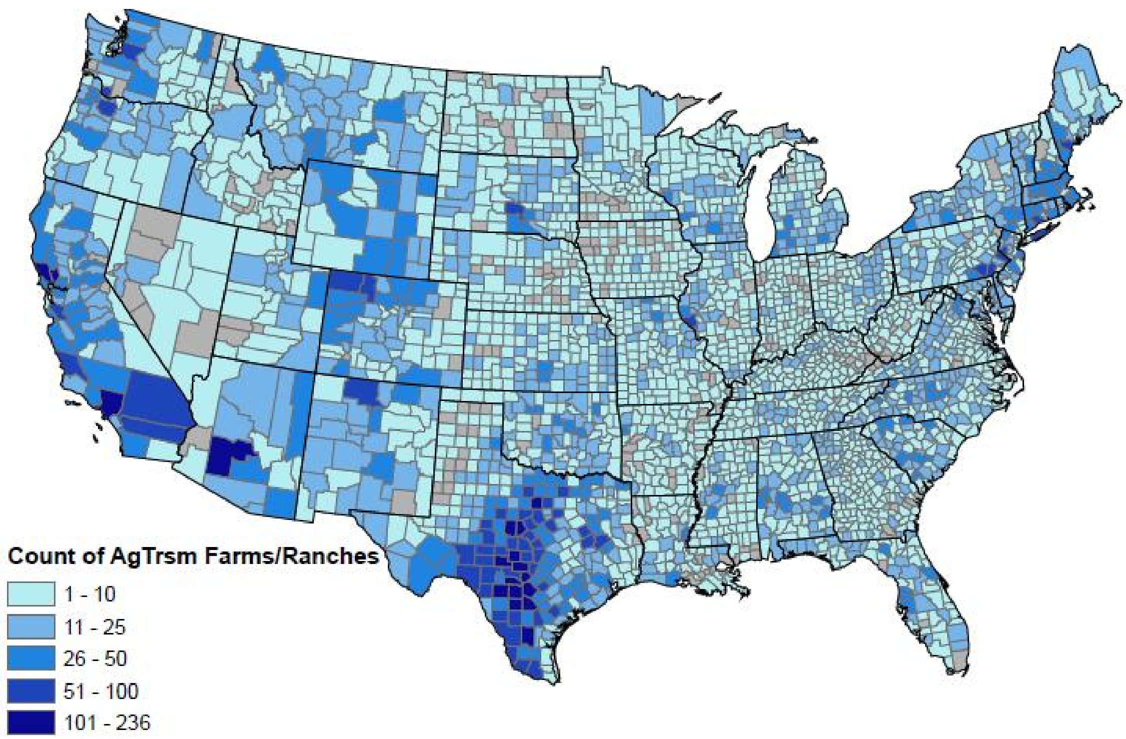


Figure 1.1 Number of Agritourism Farms/Ranches per County (2012)

As shown in table 1.1, the New England states have the highest incidence of agritourism among the contiguous 48 states. These densely populated states are an interesting juxtaposition to Wyoming and Montana, both of which are extremely rural and have the lowest population densities in the U.S., excluding Alaska. Motivations to adopt agritourism may vary across communities and farm types. For instance, the pressure to diversify in New England could be driven by the threat of development pressures related to urban sprawl, or in a positive tone, the opportunity to take advantage of the large population and traveler spending dollars in the region.

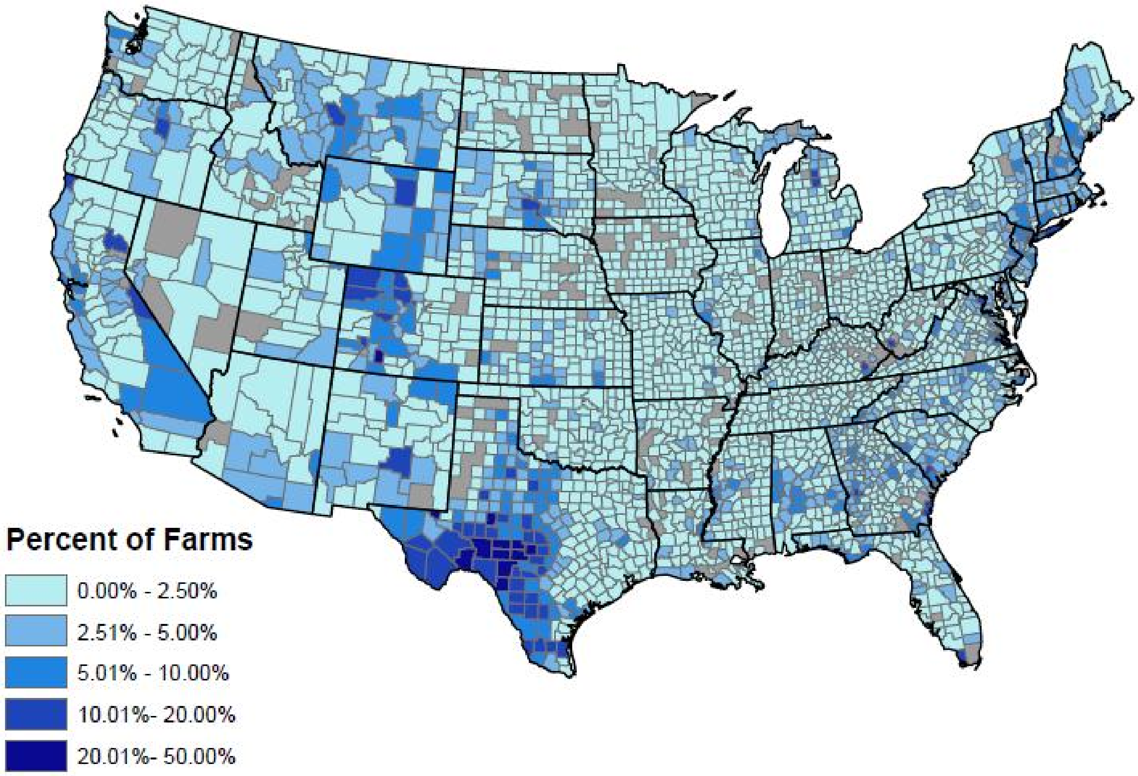


Figure 1.2 Share of Farms and Ranches with Agritourism Income (2012)

In contrast, the motivation for farmers and ranchers in the West could be more strongly related to leveraging the tourism draw that already exists in their communities due to the natural resource base and public lands of the West, and more specifically, the complementary demand for hunting, fishing, and other outdoor activities. Not surprisingly, in terms of absolute numbers, the larger agricultural states with the largest populations or higher dependence on an agricultural economic base tend to have the most agritourism farms and revenue (Van Sandt and Thilmany, 2016). Whatever the motivation to adopt, it is likely that producer goals and the success of the agritourism operation are somewhat dependent on the location and local place-based factors surrounding the agritourism farm or ranch.

Table 1.1 States with the highest agritourism farm shares, counts, and revenues							
State	AgTrsm (Share of Farms)	State	AgTrsm Farms	State	Avg. AgTrsm Rev. ($,000's)	State	Total AgTrsm Rev. ($,000's)
RI	5.47%	TX	7,775	NJ	$ 53.07	TX	$ 132,864
NH	4.33%	CA	1,699	UT	$ 46.70	CA	$ 64,520
CT	3.97%	NC	1,135	MA	$ 41.93	NY	$ 31,250
WY	3.83%	KS	1,000	CT	$ 40.07	CO	$ 28,240
NJ	3.83%	GA	944	CA	$ 37.98	GA	$ 26,044
MA	3.70%	CO	864	NY	$36.46	PA	$ 24,677
ME	3.30%	NY	857	PA	$ 33.85	MT	$ 20,310
TX	3.12%	MO	844	AZ	$ 32.73	MI	$ 18,995
MT	2.59%	OK	840	CO	$ 32.69	NJ	$ 18,416
MD	2.50%	IL	834	WY	$ 31.62	NC	$ 17,625

Agritourism is commonly referred to as an entrepreneurial activity in the literature, which leads to several implications for rural development. Rural areas have experienced lagging employment growth, specifically in manufacturing and agricultural sectors, over the past several decades contributing to higher rates of rural poverty compared to urban areas (USDA, 2018). Rural job growth decreased significantly during the great recession of 2008, and unlike urban areas, rural America has not recovered from the loss of employment opportunities. In 2015, rural employment was 4.26% lower than prerecession levels, while urban areas added 4.02% more jobs since 2008 (Weiler, 2017). Given the wide range of possible agritourism activities, multiple motivations to adopt such activities, and the potential to frame a successful enterprise in most regions of the U.S., agritourism has received attention from the United States Department of Agriculture (USDA) as a rural development tool (Bagi & Reeder, 2012a). However, in 2012 only 1.5% of farms and ranches in the U.S. reported income from agritourism, showing the potential for growth in the industry (Bagi & Reeder, 2012b).

Unlike most other agricultural diversification strategies, agritourism also has the potential to create spillover benefits for the surrounding communities, stimulating further rural economic development. When tourists visit agritourism enterprises they not only spend money at the agritourism site itself but are also likely to spend money at other local businesses such as restaurants, hotels, and shops. This infusion of money from nonresident tourists creates a relatively higher multiplier effect, indicating such activity may generate more employment opportunities throughout the surrounding local economy. For example, using IMPLAN, Thilmany, Sullins, & Ansteth (2007) estimate a net economic multiplier of 1.22 for agritourism enterprise revenues and find that the total economic contribution to Colorado from agritourism in 2006 was approximately $2.2 billion.

In contrast, other diversification strategies in the local foods economy, such as farmers markets, farm-to-table, and community supported agriculture, may not have as strong of rural-urban linkages. These strategies principally rely on reallocating food buyer demand in the rural economy to more locally marketed, owned and/or governed food firms, which may have benefits, but not to the extent of activities bringing "outside" dollars and consumers in from external buyers. This is both a benefit and a detriment to the agritourism industry. Agritourism establishments may exhibit strong urban-rural linkages through the flow of consumers passing between the site and urban area, but this also means that an agritourism business must be able to establish these rural-urban flows for the diversification strategy to be a viable option. Attracting visitors may be contingent on multiple factors including proximity to urban areas and other tourist attractions, surrounding natural amenities, agricultural production types, and the region. Due to agritourism's perceived dependence on place-based factors, agritourism is not a silver bullet in the context of farm-level and community wealth creation, and in various instances may be a second-best option relative to other alternatives in the local foods system. Research on the heterogeneity of these opportunities across places, and how it may compare or be clustered with other diversification options, is thus warranted.

Another complexity that may be a challenge to a developing agritourism industry is the lack of information on a set of best practices for business development due to different market factors underlying enterprises and clusters. While some of the other diversification strategies previously mentioned involve transporting physical and somewhat uniform goods to a marketplace, or may involve changing a production process to sell the agricultural products under a different label, agritourism enterprise development is a more intangible value-added process. Essentially, the agritourism site must develop and offer experiences that are customer service and experience oriented, and expectations and preferences of travelers may vary across regions, traveler types, and portfolio of agritourism activities. This high degree of heterogeneity in the industry presents a challenge for rural development and tourism practitioners seeking to provide pertinent business development information to agritourism operators on best practices, agritourists' willingness to pay (WTP), travel behavior, and activity preferences.

To partially address the limited information available to guide agritourism enterprise development, this dissertation presents a three-pronged approach to inform the literature on the relevant aspects on the diversity of the agritourism industry and explores potential policy implications from these empirical findings. The rest of the dissertation is broken into three chapters of applied research and discussion, each with their own literature review, theoretical model, and empirical analysis, and a final chapter synthesizing overarching conclusions. The three content chapters provide a comprehensive view of the complexities of the agritourism industry by addressing: 1) the supply of agritourism and factors that influence where firms locate; 2) the revealed traveler demand for agritourism; and 3) the relative marginal values consumers place on various agritourism attributes.

Chapter two addresses the importance of place-based factors in agritourism adoption and success. Simply mapping the number of agritourism establishments or the amount of agritourism revenue by county for the United States shows interesting patterns and clusters of activity, but very little is known about why we see agritourism where we do and what location implies for agritourism establishment revenues. Using a spatial interactions model framework, and three different regional economic theories to motivate the research questions and hypotheses, I develop an empirical model to explore spatial patterns. Using farm-level data from the 2012 Census of Agriculture to analyze factors that affect the wide range of supply factors (community and farm characteristics), this study shares relevant spatial patterns that may exist among agritourism establishments in the United States. Results highlight the importance of place-based factors in determining both the likelihood of adopting an agritourism enterprise and the expected gross revenue of an agritourism enterprise. The findings provide supporting evidence for all three trade theories tested with the data, perhaps suggesting that different economic forces are at play in guiding the adoption decision in different areas of the US. Specifically, the discussion acknowledges the important role natural endowments, operator experience, and economies of agglomeration play in the success of an agritourism farm or ranch.

Chapter three explores existing agritourism demand in the Western United States and directly addresses a bias in the travel cost method (TCM) that has confounded the literature since the method's origin. Using primary data collected through a national online survey, I develop a flexible functional form of the TCM to capture how differences in traveler behavior, regions, and agritourism activities influence the value a visitor places on the agritourism experience.

Specifically, I estimate consumer surplus values, which represent the value a consumer receives above the market price (cost of admission), in order to identify comparative advantages across regions. While the flexible functional form is unique in the literature and furthers the understanding of heterogeneity in travel cost models, perhaps the most significant contribution this essay makes is its more precise estimation of the multi-destination bias and how it may vary across different TCM studies. Beyond the significance of the multi-destination bias and its implications for future TCM studies, applied results highlight the different behaviors across traveler types and identify comparative advantages, varying by type of agritourism activities, for a number of regions.

Chapter four complements the revealed preference model in chapter three by using a stated preference model to explore how differences in consumer preferences are influenced by the consumers' location of origin, and how the WTP for specific agritourism qualities may differ depending on the location of the agritourism activity. Applying primary data from an online choice experiment, a random utility model framework is developed and estimated using a latent class logit model to identify market segments and estimate the relative marginal values consumers place on different agritourism experience characteristics. Using the model's market segment and WTP estimates, I then use a geographical interpolation tool to map consumer preferences across the contiguous United States in order to differentiate how regions may vary in their comparative advantage for potential agritourism enterprises.

This three-pronged approach provides a comprehensive analysis of a diverse industry that varies across locations, agricultural production types, and traveler types. Agritourism experiences vary depending on the type of agritourism and traveler, both of which are likely to change across space, and the analysis of the industry is further complicated by a lack of a standardized type of enterprise. This presents an interesting challenge in exploring the supply, demand, and consumer preferences for agritourism. However, by focusing on the influence of location and place-based factors and applying flexible empirical models that account for the multiple sources of heterogeneity that may be important for enterprise development, I provide valuable information for stakeholders. Moreover, broader contributions to the tourism economics literature are offered by pushing the frontier of research through development of greater precision in the data collection methods and subsequent CS estimates in the Travel Cost Model.

CHAPTER TWO: PLACE-BASED FACTORS AND THE SUCCESS OF FARM-LEVEL ENTREPRENEURSHIP

Introduction

Innovation and entrepreneurship are two central drivers of industry growth across markets, but their determinants likely differ depending on the industry in question. As the agricultural sector becomes increasingly competitive, U.S. farms and ranches, particularly those of small- to medium-scale, have sought out diversification strategies by tapping into new niche markets and developing value added products to address a variety of financial and personal goals (Bauman, Thilmany, Jablonski, & Shideler, 2015; Low et al., 2015). One such diversification strategy, agritourism, has received particular attention from agricultural operators, economic development practitioners, and policy makers alike due to its perceived flexibility to complement multiple agricultural production activities, and its potential to act as a catalyst for rural development (Skuras, Dimara, & Petrou, 2006; Thilmany et al, 2007; Das & Rainey, 2010).

However, the vast degree of heterogeneity in agritourism activities across multiple agricultural businesses in different regions makes the drivers of innovation and entrepreneurship in this industry difficult to identify. For example, the factors motivating an operator to start an agritourism enterprise may be much different for a winery operator in California than for a rancher in Wyoming. Furthermore, once developed, these agritourism enterprises will likely generate different degrees of economic activity due to differences in their agricultural production enterprises, surrounding place-based factors, and community attributes. Nonetheless, identifying the firm and locational factors that influence participation in this entrepreneurial activity may assist agricultural operators and rural development practitioners in recognizing and leveraging firm and location specific attributes to mitigate market pressures and stimulate rural development in their local economies.

Considering the unique spatial dimensions in the agritourism industry, this analysis attempts to fill a gap in the literature by exploring the following research question: How do operational characteristics and place-based factors, such as natural amenities, populations, and transportation infrastructure, relate to the propensity to adopt an agritourism enterprise? And moreover, how does the level of revenue, a proxy for economic activity, generated by that enterprise vary by these factors?

Over the past few decades, increasing market pressures, a growing public interest in food systems, and attention to new models of rural development have motivated an increasing number of agritourism studies. Most of the existing agritourism literature has focused on identifying operators' motivations to adopt, implications for rural development, or estimating the demand for various agritourism regions and activities. However, the relatively small agritourism literature falls short of exploring the complex drivers and interdependency between location and market potential to support the viability of individual or clusters of agritourism establishments.

While relatively well studied, principal operators' motivations to adopt agritourism likely vary depending on their farm or ranch's location. Some operators may be compelled to adopt an agritourism enterprise in order to take advantage of an emerging market and reduce the financial pressures from urban sprawl, while another set of more rural operators may develop an agritourism business to create economic opportunities where there otherwise were none (Nickerson, Black, & McCool, 2001; McGehee & Kim, 2004; Barbieri & Mahoney, 2009; Tew & Barbieri, 2012; Barbieri, 2013). Understanding how a producer's location influences the decision to adopt such an enterprise, and further, the expected revenues from their agritourism operation will provide researchers with a more general framework in which to view these previous case studies focused on operator motivations.

In general, the literature finds mixed results on whether agritourism can enhance a farm or ranch's profitability. In a survey of over 800 farms across the U.S., Barbieri (2013) find that agritourism increases farm profitability more compared to other diversification strategies. Schilling, Sullivan, & Komar (2016) find that while agritourism can increase the profitability of "intermediate-scale and lifestyle farms," agritourism has no noticeable effects on the profitability of larger, commercial scale farms. Considering the mixed findings of other studies (i.e. Barbieri (2013), Busby and Rendle (2000), Schilling et al. (2016), and Tew & Barbieri (2012)), it seems that the differences in agritourism experiences across agricultural production types, size, and location play a large role in whether agritourism can significantly increase farm profitability. However, the reader should keep in mind that this chapter focuses on agritourism revenue rather than profitability due to the interest in identifying the pull-factors of agritourism.

Another branch of literature that is of particular interest to rural development practitioners is the role of agritourism to act as a tool for rural economic growth. As the farmscape of the U.S. continues to evolve, it is becoming increasingly difficult for small- and medium-sized farms and ranches that cannot take advantage of economies of scale to succeed in more competitive agricultural commodity markets (Che, A. Veeck, & G. Veeck, 2005; Hoppe & MacDonald, 2013). As a tourism enterprise, agritourism can act as an export industry bringing dollars into rural communities where multiplier effects stimulate economic growth in multiple sectors (Skuras et al, 2006; Thilmany et al, 2007; Das and Rainey, 2010). Identifying place-based factors associated with increased agritourism activity may assist policy makers and rural development and tourism practitioners alike in leveraging community strengths to increase agritourism traffic in their local communities. Finally, the last main branch of literature, focused on consumer demand and willingness to pay estimates (Carpio, Wohlgenant, & Boonsaeng, 2008; Hill, Loomis, Thilmany, & Sullins, 2014; Van Sandt, Low, Thilmany, 2018), will benefit from this study's complementary supply-side analysis from which researchers can draw comparisons between producer and consumer behavior.

After conceptualizing a theoretical framework that derives from the gravity model, farm-level data from the USDA Census of Agriculture (2012) is applied to a two-stage Heckman model to empirically test several hypotheses surrounding firm location and economic activity. This approach lends to exploring the potential role of three regional trade theories (Heckscher-Ohlin, Ricardian, and agglomeration) in the context of the locational aspects of the agritourism sector while controlling for potential selection bias. The gravity model has been used extensively since the 1960's to describe international trade, and more pertinent to this study, in estimating tourism demand (Morley, Rossello, & Santana-Gallego, 2014: Keum, 2010). In the context of agritourism, the gravity model is applied to analyze the pull factors of agritourism, or more specifically, the factors that cause agritourism to generate more agritourism revenue (a proxy for relative economic activity). This method enables me to determine how natural endowments, proximity to natural and urban assets, as well as operator experience and efficiency and travel infrastructure, influences an operator's decision to adopt and explore the importance of location in the success of an agritourism establishment.

The chapter proceeds with a review of the literature on agritourism and entrepreneurship, and then presents a theoretical structure to frame the empirical analysis. Results from the empirical analysis suggest that agritourism generally benefits from tourism-oriented industry clusters and that the success of an operation is also dependent on the type of agricultural products produced on the farm or ranch, both of which serve as a potential pull to visitors. Additionally, agritourism operations may be more successful the closer they are to larger cities and natural amenities like National Parks Service (NPS) assets. Results should be of particular interest to local economic development practitioners as well as policy makers who are interested in rural development and the resilience of agricultural communities in the face of global market pressures that affect the agricultural sector (and other primary sectors) disproportionately.

Literature Review

The development of industries is marked by the continuous cycle of consolidations, new niche players, and unique value-added products. While some businesses may benefit from consolidating with larger firms through cost reductions, product branding, or more efficient distribution, other businesses may continue to grow independently through innovation or by tapping into emerging markets. In perfectly competitive markets, innovation is a central driver of short-run economic profits, and an essential factor for long-term business growth. In addition to these firm-based benefits, the greater economic activity generated through innovation leads to greater local employment growth and significantly contributes to the general economic prosperity of the local economy (Duranton, 2007).

A significant body of literature is dedicated to identifying drivers of innovation and entrepreneurship and their regional economic implications. While some innovation is born out of greater technological efficiencies or advancements, other sources of innovation and entrepreneurship include comparative advantages in natural endowments or between/within industry agglomeration. For example, while the invention of the moving assembly line reduced Ford Motor Company's automobile operating costs, allowing more vehicles to be sold at a lower price, the proximity of the automobile hub to factor endowments and forward- and backward-linkages also played a significant role in establishing Detroit as the historic automobile capital of the U.S. (Klein & Crafts, 2012). Examples of the importance of local market potential, forward- and backward-linkages, scale effects, economies of agglomeration, and factor endowments toward the growth of new and existing industries are abundant in the regional literature.

Similar to many other industries, the evolution of the agricultural sector is a story of consolidation, new niche markets, and value-added products. Between 1950 and 2012, the number of farms in the U.S. decreased by 60.86%, the average farm size doubled, and the total real value of production[1] increased 74.8% (USDA, 2014a). While much of the growth in the value of this industry is due to technological advancements in the mechanical, chemical, and biological factors of production, a set of these agricultural businesses have evolved by either entering niche markets (i.e. organic and local), or by offering value added products (i.e. salsas, jams, and leather crafts).

As is consistent with other entrepreneurs (Verheul, Thurik, Hessels, & van der Zwan, 2010), farmers and ranchers may be innovating and entering these alternative markets out of necessity or opportunity (Lasandahasi & Kodithuwakku, 2011; Low, Henderson, & Weiler, 2005). For example, a farm or ranch's proximity to urban markets, increasing land values, low ag commodity prices, and the growing public interest in differentiated food products may motivate some producers to take on direct to consumer or retailer marketing. U.S. Department of Agriculture's Economic Research Service data shows clearly that small-scale operations are more likely to have a 'high risk' operating profit margin (USDA, 2018). These producers are likely less able to take full advantage of new production technologies due to their scale, and may

[1] Total real value of production includes the value of crop and animal production as well as farm related income adjusted for inflation.

increasingly need to innovate by exploring new niche-markets and value-added products to ensure higher value production (Bauman et al, 2015; Low et al, 2015).

In addition to producers' motivations for adopting various diversification strategies, some alternative marketing efforts have received significant attention from policy makers and researchers in recent years due to their potential to create wealth in rural communities. On average, rural America has observed negative or stagnant growth in multiple measures of economic prosperity in recent years. Roughly 46 million people live in rural America, however the share of Americans living in rural areas has been decreasing significantly due to youth outmigration, fewer births, and increased working-age adult mortality (USDA, 2018). In addition to population loss, the average median income for rural counties is roughly 25% below that of urban areas, and rural employment levels have yet to return to their pre-recession levels (USDA).

Given the transformation of the agricultural sector and stagnation of economic growth in rural areas, policy makers and researchers have turned to programming that strengthens rural-urban linkages as a potential source of economic development. Building stronger partnerships between rural and urban communities may lead to greater economic growth and development in both communities, such as greater access to employment opportunities, natural amenities and resources, and economies of scale for public services (OECD, 2016). While examples of rural-urban partnerships exist throughout the world, one such example in the U.S. is the Lexington Bluegrass Alliance in Kentucky (OECD). The alliance consists of nine counties clustered around Lexington, Kentucky, and is committed to supporting greater entrepreneurship, firm growth, and business relocation (OECD).

Amongst other diversification strategies with the potential to establish rural-urban linkages, agritourism has been found to be relatively more successful at increasing farm profitability (Barbieri, 2013), particularly for intermediate-scale farms and ranches (Schilling et al, 2016), and can lead to higher per capita incomes for some regions (Brown, Goetz, Ahearn, & Liang, 2014). These spillover benefits originate from the fact that agritourism can act as an export industry by bringing dollars into rural economies and creating a multiplier effect that can stimulate other related sectors in the local rural economy (Skuras et al, 2006; Thilmany et al, 2007).

Agritourism is most commonly defined in the literature and by the USDA as any revenue generating enterprise that takes place on a working farm or ranch and provides some service or recreation to visitors for a monetary fee (Arroyo, Barbieri, & Rich, 2013). Examples of agritourism include wineries, farm stands, u-pick fruit/vegetables, farm dinners, fishing, hunting, pumpkin patches, and corn mazes just to name a few. Between 2002 and 2012, the number of farms and ranches with agritourism operations grew over 18% and real average agritourism revenue per agritourism establishment increased by 130% (USDA, 2014a). However, the share of farms and ranches with agritourism operations is still relatively small, with only 1.5% of farms and ranches operating an agritourism enterprise in 2012, indicating room for growth in the industry (Bagi & Reeder, 2012b). While agritourism operators' motivations are relatively well understood, an underexplored area is the operational and locational factors that lead to a farm or ranch operator starting an agritourism enterprise, and how these factors contribute to the economic activity generated by that enterprise.

The entrepreneurial activity in a region can be measured multiple ways, including, the number of startups, the share of non-farm proprietors in the local labor force (also called breadth), the number of patents filed, average firm size, and venture capital financing. While, somewhat dependent on the metric used, in general, regional economies with greater levels of entrepreneurship tend to grow faster and experience higher employment growth relative to their counterparts. (Glaeser, S. Kerr, & W. Kerr 2012; Glaeser, Kallal, & Scheinkman, 1992; Chatterji, Glaeser, & Kerr, 2014). Relative to other small firms, agritourism businesses may be particularly conducive to economic development simply due to the nature of tourism bringing in dollars from outside the community. Using county level data, Brown et al. (2014) analyze this economic development role of agritourism and find that while agritourism plays an insignificant role in increasing personal incomes at a national level, it did lead to increased personal income in some regions of the US. Other studies have identified significant multipliers tied to agritourism economic activity (Skuras et al, 2006; Thilmany et al., 2007) that show the full economic impact of agritourists' expenditures at an agritourism site. While it is relatively clear that agritourism can support regional economic development, previous studies have shown that these benefits are not consistent across all regions and communities, necessitating further exploration into how these entrepreneurs choose to adopt an agritourism enterprise, and what regional or operational characteristics lead to the success of that enterprise.

Identifying Sources of Comparative Advantage

Different regions may have comparative advantages in certain types of tourism when compared to other regions. There are multiple trade theories that seek to explain comparative advantages within more traditional economic industries (i.e. manufacturing, agriculture, etc.) that have also been applied to understanding the comparative advantages in the tourism industry. Zhang & Jensen (2007) outline several of these trade theories in in the context of tourism including the Heckscher-Ohlin, Ricardian, and agglomeration trade theories. These theories can be used to describe different sources of comparative advantage specific to tourism in certain regions based on natural endowments (Heckscher-Ohlin), relative productive and technological efficiency (Ricardian), and gains from the surrounding area's infrastructure and externalities (agglomeration) (Zhang and Jensen). As pointed out by Gray (1989), no single trade theory can explain every type of trade flows singularly, and accordingly, this research incorporates different aspects of the potential types of comparative advantage in agritourism from the three trade theories mentioned above.

The Heckscher-Ohlin theory of comparative advantage asserts that a region will export goods/services that are produced using factors that are relatively abundant in that region. Examples that might provide evidence of this theory in the agritourism industry include: outdoor recreation activities on farms and ranches in the Rocky Mountains, farm stands/stores near urban areas and busy transportation corridors, a microclimate that allows for unique agricultural products and agritourism offerings that are conducive to human interactions. This theory has particular importance in the tourism sector, which requires physically attracting visitors to a specific location (Carreras & Verdaguer, 1995; Urtasun & Gutierrez, 2006). Hill et al. (2014) and Carpio et al. (2008) both find more agritourism trips are taken to areas with high natural amenities, indicating that agricultural operators in these natural amenity rich areas are more likely to generate more economic activity from adopting an agritourism enterprise. Similarly, farms and ranches near population centers are more likely to adopt an agritourism enterprise, most likely due to their proximity to urban buying dollars (Bagi & Reeder, 2012; Che et al, 2005; Veeck et al, 2006).

The Ricardian theory of comparative advantage identifies a region's relative productive and technical efficiency as the primary driver of which goods/services that region exports. At the agritourism firm level, there are multiple sources of productive and technical efficiencies that one might hypothesize, including: number of years in operation, the principal operator's age, and participation in other diversification strategies. In a survey of farms and ranches in Texas, Barbieri and Mahoney (2009) find that nearly three quarters of diversified operations had at least three types of diversification out of eight categories. This is particularly interesting, as it may provide evidence of possible synergies between diversification strategies such as reduced marketing costs, shared capital requirements, and other cost reducing or revenue generating interactions.

Economies of agglomeration occur when a local industry benefits more from the addition of a new firm locating in the cluster than if that firm had located outside of the cluster. There are two sources of economies of agglomeration presented in the literature. The first, identified by Jacobs (1969), argues that economies of agglomeration occur when that additional firm is from a diverse industry, spurring experimentation and innovation through inter-industry knowledge transfers. The second, developed by Marshal (1890), Arrow (1962), and Romer (1986), but formalized by Glaeser et al. (1992), argues that the benefits from agglomeration occur when the additional firm is from the same industry creating a concentration of specialized firms who benefit from sharing similar infrastructure, labor skills, and industry specific knowledge (van der Panne, 2004).

While there is empirical evidence to support both theories in different settings and depending on the nature of the industry in question, agritourism firms likely benefit from the Marshall-Arrow-Romer theory by being located in clusters of similar firms in the tourism industry. Although not explicitly labeled "agglomeration," Che et al. (2005) find evidence of "value webs" in the Michigan agritourism industry, where individual agritourism establishments benefit from coordinating with surrounding agritourism businesses. Michigan agritourism farms benefit from these value webs in the form of information sharing, business referrals, purchasing linkages, regional reputation, and increasing the total market size to a greater extent than if they were to just try to increase their individual market share. Two other reasons agglomeration might exist in the agritourism industry are: 1) agritourism clusters are more attractive for multi-destination travelers looking to partake in multiple experiences, and 2) agritourism clusters reduce search costs for consumers (Stahl, 1982; Urtasun & Gutierrez, 2006)

Economies of agglomeration differs from the previous two theories in that although it can be a source of comparative advantage, the presence of agglomeration can also have economic development implications, particularly in the tourism sector. In a Spanish case study by Urtasun & Gutierrez (2006), tourism development was found to increase per capita income in regions with less economic development in other sectors. Furthermore, as was previously mentioned, many small establishments lead to greater employment growth relative to fewer larger firms (Chatterji et al, 2014). However, it should be noted that agglomeration economies experienced by small firms are more sensitive to distance decay due to smaller firms' greater reliance on external knowledge (van der Panne, 2004).

Methods

The gravity model was historically used to measure international trade flows driven by countries' comparative advantages. The literature also shows its successful application in estimating tourism flows – arguably, a form of trade (Morley et al, 2014; Keum, 2010). Early attempts to apply the gravity model to tourism flows were challenged due to unrealistic assumptions and a lack of a theoretical foundation (Morley et al, 2014). In the past decade, however, the linking of the gravity equation to individual utility theory and empirical support has made the gravity model a more widely accepted tool and an empirically-validated approach to assessing both international goods and tourism flows (Morley et al, 2014; Keum, 2010; Kimura

& Lee, 2006). Fourie & Santana-Gallego (2013) apply a gravity equation to observe tourism flows across 159 countries and found cultural affinity and ethnic reunion (traveling to one's ancestral region) play significant roles in explaining tourism patterns. This indicates that regional differences in agritourism activity may not only be due to the comparative advantage of that region, but also the area's associated subculture.

In order to explore the success of an agritourism enterprise in the context of the gravity model, this study applies farm-level data (1.3 million observations) from the USDA's 2012 Census of Agriculture to a two-stage Heckman model. The first stage, or the selection equation, corrects for endogeneity resulting from nonrandom self-selection bias using a probit model to estimate what factors influence the decision by a farm or ranch to adopt agritourism enterprises or not. The second stage (outcome equation) utilizes the inverse Mills ratio from the probit model in its estimation of farm-level agritourism revenue (as a proxy for the relative economic activity generated by the site).[2]

The independent variables of the gravity model can be broken down into two categories which are derived from Isaac Newton's original equation describing gravitation attraction: mass and distance (equation 2.1). For the application of the gravity model in this study, "Mass" has two different forces: push- and pull-factors. Push-factors are those causing consumers to actively seek out agritourism activities and can be thought of as factors of demand. Pull-factors represent what causes agritourism to be successful in a particular area and may include the physical features of the area, the number of clustered agritourism sites, or the types of activities offered. Herein we focus our attention on pull-factors as revealed demand models have been used for

[2] Several of the county level variables used in this modeling process were graciously generated by the GIS team at the Economic Research Service, USDA, using raster level data, while the other county level variables came from secondary sources (USDA, 2014a; USDA, 2014b; NPS, 2014).

exploring the push-factors for agritourism such as personal income, travel expenses, and preferences (Carpio et al., 2008; Hill et al., 2014). Focusing on the pull-factors offers an important contribution to the literature in that it provides producers with valuable information about the potential draw of their location and effective managerial decisions. Additionally, insight into pull-factors provides a complement to the existing consumer choice literature.

$$Measure\ of\ Attraction_i = \frac{Mass_i^{\beta}}{Distance_i^{\alpha}} \quad (2.1)$$

While motivated by the theory, applying the traditional gravity model may not be the best modeling approach since a significant share of agritourism trips are reported to be made by multi-destination travelers (Hill et al. 2014). In addition, some variables may be both push and pull factors. Subsequently, including the distance term would give an incorrect measure of friction that an agritourist may experience in visiting a site since it does not account for the additional benefits of visiting other sites that the tourist takes into account when making travel plans. Additionally, some variables such as the travel time to a town or NPS lands may be interpreted as pull factors since they bring tourists to the area of the agritourism site. However, one might consider that place-based assets may also be interpreted as push factors since agritourists may be motivated by a getaway effect to leave the congestion of cities for a more rural landscape, or alternatively, a traveler to a national park or forest may add an excursion to complement other outdoor recreation activities.

Due to the shortcomings of the traditional gravity model in the present context, this research stays true to its theoretical intuitions, while empirically estimating what could more accurately be called a "unilateral spatial interaction model." This generalized form still considers the complementarity of demand and supply, intervening opportunities from substitutes, and the transferability of travelers, preserving the spatial interactions underlying the gravity model. The empirical results may still be interpreted as key pull factors, which contribute to an agritourism site's degree of attraction independent of consumer demand (push factors) and distance traveled.

The unilateral spatial interactions model (equation 2.4) is unique from the traditional gravity model, which usually compares the attraction between two regions, in that this study only seeks to measure the attraction of random agents (agritourists) to a specific set of points (agritourism sites). Essentially there is a two-dimensional plane (the contiguous U.S.) with random particles (agritourists) traveling across it in various directions. However, these random particles are not completely random since the majority of them tend to locate in large groupings (cities) and most frequently travel between these groupings. Moreover, when they are in motion, they are bound to designated paths (travel infrastructure), each with varying degrees of activity. The purpose of this model is to distinguish what factors attract these particles to particular points that may lie on less traveled paths and in areas with notably less persistent traveler activity. In the context of agritourism, the model attempts to distinguish what factors pull visitors to agritourism operations, and what factors create resistance to this attraction.

While the traditional gravity model in tourism uses the number of visitors to an area as the measure of attraction, these data are not available at a large scale and at the level of individual farms and ranches, thus agritourism revenue per farm or ranch (**Y** in equation 2.4) will be used in its stead. This is a fitting metric because if the variation in price caused by the types of activities and regional differences can be reasonably controlled for by integrating types of crops grown and regional dummy variables, then the proposed metric can be thought of as an agritourist's "footprint."

$$Prob(\boldsymbol{A} = 1|\overline{\boldsymbol{Z}}) = \Phi(\overline{\boldsymbol{Z}}\boldsymbol{\theta}) \quad (2.2)$$

Where, $\boldsymbol{A}$ is an $nx1$ vector with entries equaling one if the agricultural business has an agritourism enterprise and zero if otherwise, $\overline{\boldsymbol{Z}}$ is an nxm matrix of independent variables related to participation in agritourism, Φ is the normal cumulative distribution function, and $\boldsymbol{\theta}$ is an $mx1$ vector of coefficients.

(2.3) $$\boldsymbol{\eta}(\boldsymbol{Z}\,\boldsymbol{\theta}) = \frac{\phi(\overline{\boldsymbol{Z}}\widehat{\boldsymbol{\theta}})}{\Phi(\overline{\boldsymbol{Z}}\widehat{\boldsymbol{\theta}})}\,, \qquad where\ \eta \sim N(\mu, \sigma^2)$$

Where, $\boldsymbol{\eta}$ is an $nx1$ vector of inverse Mills ratios and ϕ is the probit probability function.

(2.4) $$\boldsymbol{E}(\boldsymbol{Y}|\overline{\boldsymbol{X}}, \boldsymbol{A} = \boldsymbol{1}) = \boldsymbol{\alpha_0} + \sum_{k=1}^{K} \overline{\boldsymbol{X}}_k \boldsymbol{\beta}_k + \boldsymbol{\eta}\lambda + \boldsymbol{\varepsilon}$$

Where, $\boldsymbol{Y}$ is and $nx1$ vector of agritourism revenue, $\overline{\boldsymbol{X}}$ is an nxk matrix of friction, push, and pull factors, $\boldsymbol{\beta}$ is a $kx1$ vector of coefficients, λ is a $1x1$ parameter of unknown variance, and $\boldsymbol{\varepsilon}$ is an $nx1$ vector of normally distributed random error terms.

As is standard in the Heckman two-stage model, a probit model (equation 2.2) predicting a farm or ranch's participation in agritourism is first estimated to predict what influences a farm or ranch to adopt, or select into, an agritourism enterprise. Using the predicted coefficients for this model, the inverse Mills ratio is calculated for each observation (equation 2.3). Including the inverse Mills ratio in the spatial interactions model (equation 2.4) corrects for the specification bias resulting from the anticipated endogeneity of the nonrandom self-selection and the inclusion of the agglomeration variables in the outcome equation. If the coefficient of the inverse Mills ratio, λ, is significant then the $cov(\overline{\boldsymbol{X}}, \boldsymbol{\varepsilon}) \neq 0$ and the hypothesis of endogeneity is supported, justifying the use of the two-stage Heckman model.

Two other types of models that correct for the self-selection bias were considered as alternatives to the Limited Information Maximum Likelihood (LIML) two-stage Heckman model: Full Information Maximum Likelihood (FIML) and the subsample OLS (two-part model). One issue that can arise in the Heckman model occurs when there is a lack of exclusion restrictions (Little & Rubin, 1987). If this is the case, the only independent variable identifying the outcome equation is the inverse Mills ratio, which in most cases leads to issues with inefficient parameter estimates due to collinearity (Puhani, 2000). Due to this collinearity, when there are a lack of exclusion restrictions, subsample OLS is preferred to the Heckman and FIML models (Puhani). FIML models are generally preferred to LIML models such as the Heckman due to efficiency gains, however, the gains in efficiency from FIML are greatest with higher R^2s and higher correlations between the selection and outcome equations' error terms (Nelson, 1984; Puhani, 2000). Given the strong exclusion restrictions in the selection equation (total value of production and operator demographics), the large sample size and low R^2 in the estimated model, and precedence set by the literature, the efficiency gains from computationally more complex alternative models appear negligible, leading to the use of the Heckman model.

Independent variables are broken down and categorized into three of the major trade theories, which each may partially explain why agritourism operations are more successful in some locations relative to others (table 2.1). Including variables that align with the Heckscher-Ohlin, Ricardian, and agglomeration theories of trade overcomes the criticisms of Gray (1989) by not limiting the model's explanatory power for describing the place-based drivers of comparative advantage across farms and ranches. The following subsections break down the variables and their hypothesized signs based on the trade theory to which they most closely connect.

Table 2.1 Measures of Attraction

Trade Theory	Independent Variables
Heckscher-Ohlin *(Natural Endowments)*	Crops, livestock, natural amenities index*, travel time to NPS attraction*, forest products, travel time to city of ≥10k people*
Ricardian *(Productive and Technological Efficiency)*	Age, farming as primary occupation [ψ], years in operation [ψ], direct to consumer [ψ], direct to retailer [ψ], value-added products [ψ], on site packaging facility, organic certified, total value of production [ψ], entrepreneurship breadth [†]*, patents per capita [†]*
Agglomeration *(Infrastructure and Externalities)*	Byways*, interstates*, agritourism revenue per sq. farm mile* [ψ]
Demographics	Female [†], race [†], retired [†], job off farm [†], income per capita*, population*, farm-dependent*, recreation-dependent*

** Indicates that the variable is at the county level*
[†] *Indicates the variable is* only *in the selection equation*
[ψ] *Indicates the variable is* only *in the outcome equation*

Heckscher-Ohlin

Fourteen all-inclusive, yet not mutually exclusive, crop and livestock variables were included in both models to determine which types of agricultural operations are conducive to having an agritourism enterprise and to explore the intensity of economic activity catalyzed within agritourism operations. Principal operators (POs) are able to make some decisions about what crops and livestock they produce, but the broader production categories included in the model are limited to the natural endowment of various locales. Since value-added products are directly controlled for, crops and livestock that are relatively unique or ready-to-eat without processing (i.e. vegetables, fruit and nuts, grapes, specialty livestock[3], etc.) are hypothesized to lead to greater participation in agritourism and an agritourism site's level of economic activity. This hypothesis will be supported if these variables representing production types that are more

[3] Specialty livestock are defined by the USDA to include bison, deer in captivity, elk in captivity, alpacas, llamas, mink, rabbits, and other livestock not specified in the 2012 Census of Agriculture.

unique or conducive to human interaction are positive and significant in both estimation stages of the model.

Higher natural amenities and the proximity to a National Park Service (NPS) attraction are both hypothesized to encourage higher adoption rates and greater economic activity tied to agritourism; the outdoor recreation opportunities they represent may be compliments to agritourism activities. The proximity to a city of at least 10,000 people, proximity to a NPS attraction and the county population of an agritourism site are expected to lead to greater agritourism adoption and economic activity tied to agritourism. These variables represent the agricultural business' access to markets, complementary markets, and the potential market size for agritourism.

The *Travel Time to a City* variable was calculated using a road network shapefile with speed-limits in ArcGIS. Beginning and end points consisted of each county's population centroid to a city of at least 10,000 people, taking the fastest route. The population cut off for the travel time to a city of at least 10,000 people was selected after experimenting with five different population cut offs: 10k, 25k, 50k, 100k, and 250k. Each city population size was included in the model one at a time, and while they all had the same signs, the 10k population cutoff was the only significant measure of proximity to a city. This illustrates the possibility that smaller towns can be magnets for agritourism activity, which will be discussed in greater detail below.

NPS lands with less than 40,000 annual visitors were not included in the *Travel Time to a NPS Attraction* variable, since less visited NPS assets are not expected to be large enough magnets for tourists to spill over into other recreational activities in the surrounding local economy. The potential for the variable to display insignificant effects without a visitation cutoff is heightened due to the fact that NPS lands include all national monuments, forests, grasslands, and other natural assets, not just National Parks themselves. Some of NPS assets may have their designations due to their scenic beauty or cultural significance, and may not be a popular tourist attraction relative to substitute recreation destinations. Therefore, in order to capture those NPS assets with the greatest potential to create spillover agritourists, a visitation cutoff was set at 40,000 visitors per year based on a visual inspection of a histogram of all NPS lands' visitors. In order to smooth out shocks (e.g. floods, forest-fires, anniversaries, and renovations) to visitor levels, each park's annual visitor rates were averaged over five reported visitation years (2010-2014). After excluding NPS assets that did not meet the 40,000 five-year average visitation threshold, or that were simply unlikely to be agritourism magnets due to the nature and location of the attraction (e.g. monuments in Washington D.C.), 295 out of 377 NPS assets were included in the calculation of *Travel Time to a NPS Attraction*. Similar to the *Travel Time to a City of ≥10k People*, travel times were calculated using a road network shapefile with speed-limits in ArcGIS, but instead of measuring from each county's population centroid to the nearest city of at least 10,000 people, the travel times were calculated from the county's population centroid to the border of the NPS attraction.

Ricardian

Total value of production is hypothesized to be negative, reflecting the need to reallocate resources toward the agritourism enterprise. The natural log function of this variable is meant to capture the diminishing marginal value of production of resources as they are reallocated between enterprises. *Days worked off farm* and *Farming as a primary occupation* reflect the principal operator's allocation of labor between on-farm and off-farm economic endeavors, and are thus expected to display negative and positive coefficients, respectively. Finally, the operator's *Years of experience* is hypothesized to be positive and significant, reflecting the value of work experience.

Five local foods variables are included in both stages to capture any possible synergies between agritourism and these other alternative production or marketing strategies. These variables include whether the agricultural business is *Organic certified*, has an *On farm packaging facility*, produces *Value added products*, or participates in *Direct to consumer* or *Direct to retailer* marketing channels. Participating in alternative and niche markets may imply the primary operator has more experience with initiating and operating different enterprises and managing diverse revenue streams, leading them to be both more likely to adopt agritourism as well as generate more economic activity tied to the agritourism operation. However, while Barbieri and Mahoney (2009) find that 71% of diversified farms and ranches in Texas participate in at least three out of eight diversification strategies suggesting possible spillovers between enterprises, it might also be true that additional diversifying enterprises may require unique physical capital, human capital, and/or additional hours of labor that are not relevant to the agritourism enterprise.

Agglomeration

As discussed, agritourism firms likely benefits from within-industry agglomeration. This hypothesis will be supported if *Agritourism revenue per square farm mile* is positive and significant in the estimated model, implying that additional economic activity tied to agritourism in a specific agritourism site's county leads to positive externalities for that specific agritourism site. Since agglomeration can arise for multiple reasons, including natural and geographical amenities that represent competitive advantages, several variables act to control for these other sources of agglomeration including: natural amenities, region, local population size, a farm-dependent economy, a recreation-dependent economy, and the concentration of travel infrastructure in the county. While *Agritourism revenue per square farm mile* would be endogenous in a single equation model, the two-step nature of the Heckman model and the inclusion of the inverse Mills ratio preserves the independence between this measure of agglomeration and the error term in the second stage.

Empirical Results and Discussion

There are likely different classes of membership within the range of agritourism farms and ranches due to the degree of economic activity tied to their agritourism operation. For example, farms and ranches with very little agritourism revenue might simply be claiming agritourism revenue generated from more passively managed events that bring minimal customer flows such as those who stop at farm stands. While this is still technically agritourism, the pull-factors for these very small agritourism operations are likely very different compared to other, more developed agritourism operations, such as: wineries, corn-mazes, dude ranches, etc. In addition, if researchers are interested in agritourism's potential as a diversification strategy or rural development contributor, farms and ranches with very low levels of economic activity tied to their agritourism operation are most likely not of as great of interest from a policy perspective.

To overcome the issue of different classes of agritourism operations, three separate models were estimated on two different subsets of the data to provide robustness checks. The first model includes all farms and ranches, regardless of their total value of production, the second model only includes agritourism farms and ranches with more than $350,000 in total value of production, and the third model includes all farms and ranches, but uses an interaction term to identify whether agritourism businesses with at least $350,000 in total value of production and that participate in local foods (i.e. direct to consumer, direct to retailer, organic, on farm packaging facility, and value added products) behave differently than smaller farms that integrate local food activities. This size cutoff for a large farm was defined in Hoppe & MacDonald's (2013) farm typologies based on different levels of gross cash farm income for family-owned farms and ranches in the U.S. The first model likely provides the most policy relevant results as it includes all agritourism operations. However, comparing model one's results to those of two and three may provide perspective into whether or not pull-factors vary depending on the size of the agricultural businesses included in the model, as well as how the relationship between local foods operations and agritourism differs depending on the size of the agricultural business.

The reference group for model one's selection equation is: male, white, non-retired principal operators, who own a non-organic farm or ranch with no on-farm packaging facility, value-added, direct to consumer sales, or direct to retailer sales, and who are located in the West Census Region in a county that is neither farm- nor recreation-dependent. The West was chosen as the reference group due to its significant agritourism activity relative to the other regions. The outcome (revenue) equation's reference group are POs whose primary occupation is not farming, who do not produce organic, value-added (VA), or forest products, do not participate in direct to consumer or retail markets, and who own a farm or ranch in the West Census region. The results in the first stage (table 2.2) should be interpreted as the change in the z-score for a one-unit change in the independent variable, however, most of this study's primary research questions revolve around the second stage results (table 2.3) which are more straightforward and can be interpreted as the change in agritourism revenue measured in dollars from a one-unit change in the independent variable.

The rest of the empirical results section will be outlined by first focusing on the motivators to adopt in the selection equations, and then on the pull-factors of agritourism in the outcome equations. The sets of variables in each equation are organized by data type in tables 2.2 and 2.3, but are discussed in the context of the theory of comparative advantage to which they most closely relate in following paragraphs. Most of the results and discussion to follow focus on model one's results, however some inferences are made by comparing the three models when it helps inform the discussion. Model three will only be addressed in the discussion of the second stage results since model three has the exact same first stage specification as model one.

Table 2.2 Heckman Selection Equation – Probit model
Dependent variable: 1 if Agritourism, 0 otherwise

	Variable	Coefficients		
		Model 1	Model 2	Model 3
Farm level	Intercept	-2.2416***	-2.1237***	-2.2416***
	Female	0.0402***	0.1652***	0.0402***
	Black	-0.0833***	-0.0357	-0.0833***
	Asian	0.0748*	-0.0110	0.0748*
	Hawaiian	0.1666	0.2279	0.1666
	American Indian	-0.1700***	-0.0128	-0.1700***
	Retired	-0.1919***	-0.0348	-0.1919***
	Age	0.0038***	0.0046***	0.0038***
	Days worked off farm	-0.0427***	-0.0070	-0.0427***
	Crops and Livestock			
Farm level	Hay and grains	0.0379***	-0.0053	0.0379***
	Christmas trees	0.3964***	0.4440***	0.3964***
	Maple products	-0.0825**	-0.2212*	-0.0825**
	Bee products	0.0887***	0.1879***	0.0887***
	Vegetables	0.2609***	0.1697***	0.2609***
	Fruit and Nuts	0.2663***	0.1832***	0.2663***
	Berries	0.1746***	0.3433***	0.1746***
	Grapes	0.4697***	0.3418***	0.4697***
	Cattle	0.0665***	0.1472***	0.0665***
	Horses	0.1526***	0.3120***	0.1526***
	Sheep and goats	0.1848***	0.2059***	0.1848***
	Pigs	0.0516***	-0.1249***	0.0516***
	Poultry	-0.0489***	-0.2236***	-0.0489***
	Other livestock	0.4130***	0.4961***	0.4130***
	Forest products	0.3197***	0.5663***	0.3197***

	Table 2.2 continued….			
	Value added products	0.2302***	0.3515***	0.2302***
	Direct to consumer	0.1497***	0.3860***	0.1497***
	Direct to retailer	0.3312***	0.3265***	0.3312***
	Organic certified	0.0340	-0.0278	0.0340
	On-farm packaging facility	0.0607***	0.0837*	0.0607***
Spatial				
County level	South	0.0733***	0.0004	0.0733***
	Midwest	-0.2282***	-0.3788***	-0.2282***
	Northeast	0.0704***	-0.0546*	0.0704***
	Miles of Byways/100 sq. mi.	0.0172***	0.0176***	0.0172***
	(Miles of Byways/100 sq. mi.)^2	-0.0008***	-0.0006*	-0.0008***
	Miles of interstates/100 sq. mi.	-0.0025	0.0038	-0.0025
	(Miles of interstates/100 sq. mi.)^2	-0.0006	-0.0007	-0.0006
	Ln(population)	-0.0442***	-0.0636***	-0.0442***
	Farm-dependent	0.0868***	0.0920***	0.0868***
	Recreation-dependent	0.0377***	0.0246	0.0377***
	Entrepreneurship			
	Breadth	0.6051***	0.7265***	0.6051***
	Patents per 1,000 people	0.0062***	0.0112***	0.0062***

* Significant at 1% level
** Significant at 5% level
*** Significant at 10% level

Table 2.3 Heckman Outcome Equation – Ordinary Least Squares
Dependent variable: Agritourism Revenue (Dollars)

	Variable	Coefficients Model 1	Model 2	Model 3
Farm level	Intercept	28,947.00	-329.54	27,735.00
	Total value of production	0.0012**	0.0010	0.0012**
	Retired	-7,992.43***	-14,534.00	-8,027.72***
	Acres	2.59***	3.60***	2.58***
	Age	-122.49	315.50	-119.54
	Farming as primary occupation	394.36	-23,968.00***	345.70
	Years in operation	131.09**	-119.45	133.18**
Crops and Livestock				
Farm level	Hay and grains	-1,330.04	13,680.00*	-1,250.82
	Christmas trees	-8,850.48	21,639.00	-9,087.68
	Maple products	-10,093.00	-2,462.92	-9,404.78
	Bees	-10,165.00***	-6,236.70	-10,069.00***
	Vegetables	3,604.44	30,636.00**	3,771.58
	Fruit and Nuts	24,458.00***	40,456.00***	23,928.00***
	Berries	-6,481.13*	15,457.00	-5,525.87
	Grapes	58,208.00***	133,681.00***	56,447.00***

	Table 2.3 continued…			
	Cattle	-9,281.75***	-31,967.00***	-9,163.53***
	Horses	2,706.40	4,712.59	2,822.73
	Sheep and goats	-4,736.58*	-4,056.10	-4,695.30*
	Pigs	5,326.93	4,259.80	5,706.95*
	Poultry	-2445.70	15,291.00	-2,571.90
	Specialty livestock	21,054.00***	153,442.00***	20,926.00***
	Forest products	-7,636.37**	-14,876.00	-7,856.01**
	Large Local	--	--	15,419.00***
	Value added products	9,732.77***	8,555.47	8,059.10**
	Table 2.3 continued…			
	Direct to consumer	-11,726.00***	3,119.71	-23,487.00***
	Direct to retailer	3,776.75	-24,281.00	-1,271.58
	Organic certified	-19,911.00***	-26,328.00	-22,797.00***
	On-farm packaging facility	1,000.12	6,835.80	1,218.85
Spatial				
County level	South	150.73	-2,623.01	536.30
	Midwest	6,426.78	-10,870.00	6,554.25
	Northeast	11,897.00***	-34,818.00**	12,152.00***
	Natural amenities scale	1,076.57***	-2,451.33	1,080.61**
	Minutes to population of ≥ 10,000	3.58	-128.34	3.74
	Hours to NPS attraction	-2,149.99**	-280.05	-2,091.35**
	Miles of Byways/100 sq. mi.	-475.76	616.16	-458.55
	(Miles of Byways/100 sq. mi.)^2	-8.49	-58.13	-9.79
	Miles of interstates/100 sq. mi.	-310.56	-4,291.66	-328.53
	(Miles of interstates/100 sq. mi.)^2	-22.83	622.36	-22.67
	Agritourism revenue per square farm mile	0.11***	12.11***	0.11***
	Per capita income	0.19**	-0.62**	0.19**
	Ln(population)	2,325.97***	2,463.25	2,264.38***
	Farm-dependent	-2,767.98	-9,159.77	-2,652.03
	Recreation-dependent	804.31	20,207.00**	757.78
Inverse Mills ratio		-13,422.00	15,210.00	-13,116.00
Adjusted R^2:		0.0809	0.2680	0.0814
* Significant at 1% level				
** Significant at 5% level				
*** Significant at 10% level				

Propensity to Adopt Agritourism

The Ricardian theory of trade posits comparative advantages arise out of a farm operator’s core production enterprise, experience, and specialization. Across all three models –

independent of agritourism enterprise size – older operators were consistently more likely to adopt an agritourism enterprise. However, retired principle operators were less likely to participate in agritourism, so these variables may offset one another to some degree and the status an operator reports (retired vs. actively managing) matters greatly. Rounding out operator demographics, all three models support previous findings that female principle operators are more likely to adopt an agritourism enterprise.

Several types of production specialization lead an operator to a greater propensity to adopt an agritourism enterprise, however, some production types have varying levels of impact depending on the classes of agritourism establishments included in each of the models. For example, *Cattle* has a significant and positive impact on the decision to adopt an agritourism enterprise, but the z-score coefficient is over twice as large for larger farms and ranches (model two), suggesting larger livestock operations are more likely to see value in diversifying the incomes they generate from their key assets (lands and herds/flocks). Similarly, *Horses* have a greater impact on the propensity to adopt an agritourism enterprise, particularly among higher value agricultural operations (model two). These differences in coefficients depending on the different samples included support the decision to run the model on sample subsets based on farm size classifications and interpret the nuanced differences across broader and narrower classifications of ag firms. *Horses* probably has a larger positive impact in model two because they are an integral part of more expensive agritourism activities like dude ranches.

Farms and ranches in the Midwest have a lower propensity to adopt an agritourism site relative to the West in both sample subsets. In model one, Northeast farms and ranches have a higher likelihood of participating in agritourism reflecting the relatively more well-established promotional programs, agritourism collaboratives, and community organizations when compared to the West. However, model two indicates that Western commercial farms and ranches were more likely to adopt agritourism enterprises perhaps reflecting differences in goals, agricultural production, and culture. In other words, the Northeast may have a higher share of farms and ranches with agritourism, but the West dominates all regions when only considering those agritourism operations with the highest total values of production (model two). This may be evidence of the Heckscher-Ohlin theory, where based on some natural endowment, such as climate, community assets, or culture, some farms and ranches are more likely to offer, or rather, be able to offer certain types of agritourism activities that are more lucrative than others.

Value-added, direct to consumer sales, and direct to retailer sales all lead to a greater propensity to adopt an agritourism enterprise. This may be a reflection of farms and ranches with direct to consumer sales and value-added products perceiving complementarities when adopting an agritourism enterprise. Value-added products can easily be integrated into an agritourism operation simply by selling the goods through a farm stand, whereas direct to consumer sales and direct to retailer involve interacting with individuals beyond the farm, thus providing word-of-mouth promotion that may draw visitors to the farms they purchase from at markets. More lucrative agritourism activities may be experience- rather than product-based, so they may not contribute as significantly to the agritourism revenue generated. This will be explored further in the sections to follow.

The natural log of population (*Ln(population)*) is significant and has a similar coefficient in both models, indicating that, regardless of the thresholds for agritourism sites' level of total value of production, farms and ranches in more rural areas are more likely to adopt agritourism enterprises. Similarly, farms and ranches in *Farm-* or *Recreation-dependent* counties and those in counties with scenic byways seem to have greater propensities to adopt an agritourism enterprise.

Along with the regional variables, this shows how some locations may offer more possibilities to adopt an agritourism enterprise, and in general, be more conducive to supporting agritourism activities based on the market size potential driven by surrounding population, relevant sectors, and travel infrastructure.

Finally, it appears that the overall entrepreneurial spirit of a county is a significant motivator in a farm's decision to adopt an agritourism enterprise, suggesting network or cohort effects. Both the breadth of proprietors in the labor force and the patents per capita are positive and significant at the one percent level in both models indicating the robustness of this result. These signs and the significance of the entrepreneurship variables mirror McGranahan, Wojan, & Lambert (2010) who found that both the entrepreneurial spirit, the creative class, and the combination of the two, had a significant and positive impact on job and establishment growth in rural counties for a broad set of sectors. The findings in table 2.2 imply that these synergies created from entrepreneurial activity are not only significant for general job and establishment growth, as found in McGranahan et al., but also exist specifically in the agritourism sector. The suggestion that regions with high entrepreneurship rates are more efficient in developing local agritourism industries is consistent with the Ricardian theory which posits that efficiency and specialization are the primary contributors to the comparative advantages faced by operators in this sector. The result is particularly inspiring, given that many rural communities have been forced to look for economic opportunities beyond their comparative advantages in land and labor to drive their local economies. Entrepreneurial ventures, and the evidence found here of a snow-ball effect as like-minded community networks strengthen entrepreneurial enterprises like agritourism, play particularly important roles in transitioning rural economies due to their ability to generate long-term employment opportunities.

Pull-factors of Agritourism Establishments

Thus far in the analysis, most of the discussion has focused on the relevance of the three trade theories in the context of operating an agritourism business. While this analysis is important to more completely understanding the factors that affect the decision to adopt an agritourism enterprise, the primary focus of this paper is to explore the pull-factors of agritourism. This section considers the results from the outcome equations to analyze how the three theories of comparative advantage apply to the success of an agritourism business in terms of the agritourism site's level of economic activity tied to agritourism. Models one and three reach similar conclusions given their subtle differences in model specification, however model two varies significantly due to its focus solely on larger commercial agricultural operations.

Table 2.4 Summary of Second Stage Significant Results	
Trade Theory	Independent Variables
Heckscher-Ohlin *(Natural Endowments)*	+ Crops: fruit & nuts[123], grapes[123], vegetables[2] – Crops: forest products[13] + Livestock: specialty livestock[13] – Livestock: bees[13], cattle[123] + Natural amenities[13] – Hours to NPS attraction[13] Regions (relative to West): + Northeast[123*]
Ricardian *(Productive and Technological Efficiency)*	+ Years in operation[13] – Retired[13] – Farming as a primary occupation[2] + Value added products[13] – Direct to consumer sales[13] – Organic[13]
Agglomeration *(Infrastructure and Externalities)*	+ Agritourism revenue per sq. farm mile[123] + Recreation-dependent[2]
[1], [2], and [3] indicate the variable was significant at the 5% level in models one, two, and/or three, respectively **Northeast* displayed a negative coefficient in model two	

This research elucidates Bagi & Reeder's (2012b) finding that older principal operators are more likely to have an agritourism operation by including *Retired* and *Years in operation* along with *Age* into the outcome equation. Experienced principal operators who are not involved with agritourism solely for lifestyle purposes during retirement are more likely to earn higher agritourism revenues. This result should encourage not necessarily older, but experienced principal operators to explore diversification strategies if they have the requisite expertise or complementary human capital (perhaps younger family members seeking a role in the farm's management) to integrate these new activities into their agricultural operation.

The unexpected negative signs on the local foods variables' coefficients, which indicate that other diversification strategies are substitutes rather than complements to agritourism, motivated the decision to include an interaction term between a dummy variable for large operations with the local foods dummy variables.[4] Despite *Large Local* being highly significant, it appears the signs, significance, and general magnitude of the other coefficients did not change significantly, highlighting the robustness of model one's parameter estimates. The inclusion of *Large Local* shows that larger operations actually benefit from positive spillovers between some diversification strategies and agritourism. This may be due to the different set of resources required to develop both an agritourism experience and the logistics to prepare and deliver agricultural products directly to market being too great for smaller operations to operate efficiently, so instead, evidence shows they may want to make a choice between one diversification strategy or the other. In contrast, larger operations that have more labor and access to capital may have greater flexibility to take advantage of the synergies between agritourism and other diversification strategies, and see it as a way to diversify the cash flows needed to sustain their larger operations.

[4] Large was defined as farms and ranches with a total value of production over $350,000/year. The *Large Local* interaction term took on a one if the agritourism farm or ranch was large and participated in at least one of the local foods variables: *Value added products, Direct to consumer, Direct to retailer, Organic certified,* or *On-farm packaging facility.*

Across all three models, it appears that the Heckscher-Ohlin theory of comparative advantage drives the economic success of agritourism businesses. Agricultural operations that indicated their primary production enterprise was either grapes or fruit and nuts are consistently correlated with higher levels of economic activity tied to agritourism in each model. Unsurprisingly, for agritourism operations with higher total values of production, grapes are the second leading driver in agritourism revenue, presumably because they are commonly linked to wineries – a high-value system of production directly tied to the agritourism activity of winemaking, tasting, and tours. While most of the livestock coefficients are either negative or insignificant in model one, specialty livestock is an important contributor to agritourism revenue for agritourism sites with higher total values of production, perhaps because they could contribute to heritage, outdoor recreation, or educational activities. To give an example, Frosch, Anderson, & Outlaw (2008) estimated the direct impact from deer farm hunters in Texas (considered specialty livestock), excluding all other operations, to be $73.2 million. Specialty livestock is also a significant pull-factor when including the lower value farms and ranches (model one) which may be due to these agritourism farms and ranches tapping into consumer curiosity and offering activities, such as in petting zoos with alpacas, llamas or other exotics, where the public can interact with or learn about unique animals.

The uniqueness of grapes, fruit and nuts, and specialty livestock may be what is appealing to the curiosity of travelers. Additionally, crops and livestock that are more conducive to human interactions and activities for agritourists (fruit orchards and vineyards) lead to greater revenues compared to those that produce value-added products or are more difficult to incorporate directly into an agritourism activity (hay, bees, etc.). Furthermore, it should be noted that the types of crops and livestock that were associated with a higher propensity to adopt an agritourism site also tend to generate the most agritourism revenue. This may imply that many producers are aware of the appeal their agricultural operation has with potential agritourists and are playing into their comparative advantages in agritourism activities.

Ceteris paribus, agritourism enterprises with total values of production below $350,000 in the Northeast generate more agritourism revenue. Relative to the other regions analyzed in this study, the Northeast has the most well-established state networks supporting agritourism operations (Veeck et al., 2006). This greater agritourism support in the Northeast is presumably due to the greater population density of the region necessitating agricultural operators to seek entrepreneurial responses to growing market pressures and opportunities sooner than other U.S. regions (e.g. urban sprawl, market size, etc.). Northeast agritourism establishments that are able to take advantage of economies of size and scale earn less agritourism revenue than their Western counterparts.

The natural log of county population is positive and significant in models one and three, suggesting that while small- to medium-sized agritourism establishments in less populated areas are more likely to adopt agritourism, those in more populated counties earn more revenue. This may be due to stronger perceived marketing opportunities for farms and ranches in more populated areas, whereas those in more rural counties may be adopting agritourism due to a perception of relatively few other viable diversification strategies. One pull-factor revealed in the model that could help offset the lower agritourism revenues for farms and ranches in rural areas is if they perceive and leverage access to nearby natural amenities.

Hours to a NPS attraction is significant at the 5% level when considering all farms and ranches regardless of the total value of production (models one and three). On average, an hour *closer* to a NPS attraction translates to $2,150 more in agritourism revenue. This supports the narrative that agritourism may be a complement to other types of outdoor recreation. Whether the NPS assets are pulling greater numbers of agritourists to the area, or whether they are pushing visitors out from their attractions to explore activities to supplement their NPS experience, it appears there are opportunities for co-promotion and development partnerships between the NPS and agritourism establishments.

While the higher agritourism revenues associated with the proximity to a NPS attraction could be described as a type of between-industry agglomeration described by Jacobs (1969), it appears agritourism also benefits from within-industry agglomeration (MAR theory of agglomeration). Whether it is due to information sharing, business referrals, purchasing linkages, or regional reputation, all three models provide evidence supporting Che et al.'s (2005) findings of economies of agglomeration among Michigan agritourism establishments. When including all agritourism establishments (models one and three), the expected gains for an individual agritourism establishment from agglomeration are $0.11/dollar of total agritourism revenue in the county. These gains increase dramatically to $12.11/dollar of agritourism revenues for agritourism farms with over $350,000 in total value of production. These huge benefits from agglomeration in model two are less likely due to information sharing and more likely a result of a well-established regional reputation, such as the renowned wine regions previously discussed. The other potential sources of agglomeration revolving around travel infrastructure were mostly insignificant in the second stage results.

Conclusions and Future Research

Much of the past research in agritourism has focused on the motivations of principal operators to adopt such an enterprise, the economic impact of agritourism to surrounding communities, or the factors that lead to tourists' participation in agritourism. This research expands on the literature by exploring the spatial and operation heterogeneity of the agritourism industry. Specifically, this study investigates how natural amenities, populations, infrastructure, and farm/ranch characteristics affect the decision of an agricultural operator to adopt an agritourism enterprise, as well as whether or not these factors can lead to comparative advantages in agritourism. The Heckscher-Ohlin, Ricardian, and agglomeration theories of comparative advantage are explored by applying farm-level data to a unilateral spatial interactions model framework – a derivate of the gravity model. Empirical findings suggest all three theories provide some, but varying degrees of evidence inferring how operational characteristics and place-based factors lend to comparative advantages. One limitation of the analysis is that, in the context of the spatial interactions model, the method concentrates primarily on pull-factors of agritourism activity, leaving the motivating factors for agritourists to participate – push factors – to future research. In addition, some variables in the model may be classified as both push- and pull-factors, perhaps affecting the interpretation of some results.

While all three trade theories had relevant factors that were found to be significant, the Heckscher-Ohlin theory provided the most evidence for sources of comparative advantages in regard to agritourism farms. Natural endowments, such as natural amenities, existing crops and livestock, and population, were all significant drivers in determining an agritourism site's level of economic activity tied to agritourism. With the rising pressures on small- and medium-sized farms and ranches from urban sprawl and more competitive global agricultural commodity markets, this implies producers near natural or urban amenities may be able to leverage their location to attract more agritourists and realize greater economic opportunities.

The advantages from productive and technological efficiency described by the Ricardian theory seem to be most prevalent in the adoption of agritourism rather than in actually acting as a pull-factor to grow revenues. While, older principal operators are more likely to start an agritourism business, experience running an agricultural business is more important in terms of an agritourism enterprise generating more economic activity. Similarly, while value-added, direct to consumer sales, direct to retailer sales, and on-site packaging facilities corresponded to a greater likelihood of operating an agritourism enterprise, these factors were mostly insignificant and negatively associated with agritourism revenue. It appears these other alternative marketing strategies may be complements to agritourism operations, but reduce the efficiency of smaller agricultural operations operating an agritourism enterprise. However, it appears there are synergies between at least some alternative marketing strategies and agritourism revenues reported for larger agricultural operations.

Crops and livestock may be seen as natural endowments since they rely on specific climates and geographical characteristics such as soil and terrain. For this reason, crops and livestock and locational factors were associated with the Heckscher-Ohlin theory of comparative advantage. Controlling for VA products, certain crops and livestock which may be seen as more unique (i.e. grapes, fruit and nuts, and specialty livestock) are more conducive to an operation's agritourism activity, while crops and livestock that are more common, or may be hard to convert into novel activities, decreases the site's level of attraction. This may be encouraging for farms and ranches producing specialty crops or livestock and looking for opportunities to create viable agritourism sites, especially in areas with fewer natural amenities and a prevalence of more conventional agriculture, such as in the Midwest where unique production types or methods are more likely to stand out to consumers.

Finally, although the travel infrastructure variables exhibited no or even negative effects on an agritourism site's level of attraction, positive externalities from industry concentration were significant. This evidence of agglomeration economies is of relevance to policy makers as well as tourism and economic development practitioners, but further research is needed to understand how these potential agglomeration economies can be translated to tourism development strategies. The significantly higher industry concentration benefits in the second model focusing on agritourism sites with greater total values of production venture provides some evidence of different levels of economies of agglomeration throughout the agritourism industry. However, further research should be conducted to explore how these benefits of agglomeration vary depending on the flavor of agritourism, and how these benefits of agglomeration may relate to positive spillovers for local rural communities.

Given the increasing share of farms and ranches with agritourism enterprises and the increasing market pressures on small- and medium-sized farms and ranches, understanding the role location plays in the propensity to adopt, and the success of an agritourism operation is of increasing importance. In order to fully understand the spatial relationships of agritourism and its potential role as a rural development driver, future research should explore 1) the interplay between pull- and push-factors that motivate an agritourist to participate; 2) case studies on regions with unexpectedly low/high agritourism revenues relative to their comparative advantages to identify best practices, keys to success, or barriers to growth; 3) the spatial dimensions of possible intra- or inter-industry spillover benefits by integrating NPS activity into at least one of the case studies and tourist demand studies; and 4) how the benefits of agglomeration vary depending on the flavor of agritourism and how these benefits of agglomeration may relate to positive spillovers for other local industries in rural communities.

CHAPTER THREE: EXAMINING MULTI-DESTINATION BIAS IN HETEROGENEOUS AGRITOURISM-BASED TRAVEL COST MODELS

Introduction

Since its origins in the 1970's, the travel cost method (TCM) has been a major tool in estimating consumer surplus (CS), recreation demand, total use-values, and in evaluating policies for nonmarket goods. The primary reasons for the TCM's frequent use are its ability to determine consumer surplus estimates using observed market behavior that can be collected through primary or secondary sources, and its ability to circumvent biases that can arise in stated preference models. A seemingly exhaustive set of topics from whale watching to the Grand Canyon, and deer hunting to the Great Barrier Reef have been evaluated using the TCM in order to guide policy and management decisions revolving around resource management, consumer welfare, and economic development. Despite all of the applications and developments using the TCM, one issue that has continued to challenge the literature, and has possibly misled many policy decisions, still remains: the multi-destination bias.

The central idea of the TCM is that a consumer's true willingness to pay (WTP) for a specific activity is not only the price explicitly paid to partake in the activity, but also the implicit cost of traveling to the site and the opportunity cost of time. Ceteris paribus, a rational utility maximizing consumer will travel to closer sites that minimize their costs of travel with greater frequency compared to those farther away. Given information inferred from the quantity of trips an individual takes and the associated explicit and implicit costs, the TCM allows the researcher to estimate a demand curve and calculate the consumer surplus or total value of an activity. Typically, the price variable, travel cost, is calculated by measuring the distance traveled from the visitor's point of origin to the recreation site. However, a key empirical issue arises when a traveler visits multiple destinations on a single trip and the researcher is unable to precisely determine the incremental cost in traveling to the site of interest. On a multi-destination (MD) trip, if the activity to be valued is not the primary destination, then the visitor's travel cost is no longer the distance from their point of origin to the site, but rather the incremental distance traveled from their primary path of travel to the additional destination. As may be expected, this issue arises in the vast majority of travel cost studies due to consumers commonly bundling destinations and activities to maximize their utility per dollar (and nonpecuniary time) cost of travel. Moreover, the direction and degree of the bias most likely varies by the context of the travel and sites of interest.

The key research question addressed is: can the bias in consumer surplus estimates from MD trips be reduced with more refined data collection and modeling? And, a complementary research question is: what types of destinations, activities, and contexts may be most important to address with MD trips? To answer this, a more general TCM modelling approach is developed to deal with the challenges that MD trips posed to researchers seeking CS estimates in previous travel cost models. Rather than focusing on valuing one specific site, the model is used to estimate the CS visitors receive from the broad set of agritourism sites in the Western U.S., allowing for greater insight into how MD bias can vary depending on the activity and region of interest. Agritourism is an ideal industry to use in this analysis due to its large degree of heterogeneity in activities (i.e. wine tours, corn mazes, hunting, etc.), traveler types (MD versus primary purpose travelers), and region (climate, topography, culture, etc.). The MD bias likely varies across space, travelers, and the type of recreation in the broader literature as well, so studying the bias through the lens of a heterogeneous industry is necessary to understand the nuances of the bias. Results indicate that the MD effect can lead to a significant upward or downward bias in CS estimates, depending on the geographic region and recreation experiences offered. Averaging across regions and activities, the MD bias ranges from +244% for multi-destination visitors to -10% for primary purpose visitors. Depending on the data, the magnitude of this bias could severely distort the values estimated for nonmarket goods leading to misinformed and inefficient policy and management decisions.

The possibility for biases in CS estimates from the presence of multi-destination travelers was first fully articulated by Haspel & Johnson (1982) and many studies have noted or attempted to account for multi-destination bias since the genesis of the travel cost method. Thus far, there is still no clear consensus on how to overcome the issue, or even what the consequences are from ignoring it. Some of the proposed methods to account for the multi-destination bias include dropping all MD visitors from the sample (Smith & Kopp, 1980), treating MD trips as bundles of goods (Mendelsohn et al, 1992), and more recently, using simple econometric specifications to account for shifts and rotations of the demand curve for primary purpose (PP) and MD visitors (Parsons & Wilson, 1997; Loomis,Yorizane, & Larson, 2000; Hill et al, 2014). While some papers such as Loomis et al. (2000) have attempted to estimate the bias by comparing CS estimates from multiple specifications, no one has directly estimated the bias in a travel cost model or explored how the bias may change across recreational activities and regions.

This gap in the literature has left empirical researchers attempting to correct for a bias that is not fully understood. Developing new methods to achieve greater precision in recreational demand estimates will not only save the time and resources researchers might expend using more complicated approaches, but may also provide more accurate information to policy makers who may rely on these demand estimates as a means to infer the public's value placed on nonmarket assets. In addition, comprehensively correcting for the MD bias will increase the validity of future travel cost models seeking to explore other specific biases arising from further misspecifications, such as omitting substitute prices or the inclusion of opportunity costs of time (Rosenthal, 1987; Bockstael, Strand, & Hanemann, 1987).

Rather than relying on previous methods for handling multi-destination visitors, this study makes use of a rich visitor survey dataset with questions specific to comprehensive[5] and incremental travel costs incurred, site attributes, and visitor types meticulously in order to account for the multi-destination bias. This detailed data is then applied to a flexible functional form similar to that used by Parsons & Wilson (1997) and Loomis et al. (2000) to estimate demand for agritourism across six regions, four activities, and two different traveler types. A four-stage model build-up and a unique zonal hybrid travel cost method approach illustrates the importance of accounting for industry heterogeneity and use these flexible models to identify specific trip types where correcting for multi-destination bias may be more important.

In addition to the methodological contributions made to the literature in terms of capturing site, traveler, and regional heterogeneity and the treatment of multi-destination travelers, the empirical findings are also relevant to policy makers and industry stakeholders. Due to the generalized form of the empirical model, parameters can be manipulated to create part-worth CS or elasticities of demand estimates for multiple activities and regions.

The paper progresses by first offering a closer inspection of the literature's prior handling of multi-destination biases and the strengths and weaknesses of various approaches, before proceeding to an outline of the survey instrument and the basic microeconomic theory of utility maximization used to develop a trip's Marshallian demand curve. The theoretical implications of

[5] From hence forth, "comprehensive travel costs" is defined as the travel cost calculated from a traveler's residence of origin to the recreational activity in question.

not properly measuring multi-destination travel costs are then discussed with an outline of how the empirical model more accurately accounts for biases and different sources of heterogeneity. Finally, the paper concludes with a discussion on the importance of accounting for heterogeneity, and use the results from the flexible form developed in the prior section to illustrate circumstances that may cause some travel cost studies to suffer more or less from the multi-destination bias.

Literature Review

The first paper to focus on the issue of multi-destination bias was Haspel & Johnson (1982) in their application of the TCM to Bryce Canyon National Park and its neighboring parks. In their paper, Haspel & Johnson suggested treating the average distance between visitors' destinations as the marginal travel cost to the site for any multi-destination visitors. While this approach may be an improvement from previous methods, their proxy for the marginal travel cost is really only a valid assumption if each of the destinations and the traveler's origin are equally spaced. In fact, the three parks considered in Haspel & Johnson's model are each roughly two hours apart from one another. Haspel & Johnson find no significant differences between their proposed travel cost model's CS estimates and direct measures of consumer welfare; however, Bishop & Heberlein (1979) find different consumer welfare results when comparing alternative travel cost specifications to direct measures. This may be an indication of Haspel & Johnson's (1982) relative success, or evidence that inconsistencies in parameter estimates from the MD bias might vary depending on the nature of the recreational activity.

Mendelsohn et al. (1992) took a different approach by estimating a system of demand equations for the site in question and all other possible complementary destinations or bundles of destinations. This approach requires additional assumptions to limit the number of demand equations to something reasonable for empirical estimation. As noted by Loomis (2006), the complexity of this model only makes it a practical option when the set of complement and substitute sites is relatively small.

Thus far, Parsons & Wilson (1997) offer the simplest approach to dealing with multi-destination bias. Instead of using average travel costs or a complex system of demand equations, Parsons and Wilson include the total travel costs of all visitor types in one travel cost variable and then tease out the bias by including an intercept shifter and slope rotator for multi-destination travelers. The only extra piece of information the researcher needs to know is which traveler is a PP visitor and which is a MD visitor: a fairly simple question to ask, so the ease of estimation in this specification has made it a popular approach.

Several extensions have been made to the Parsons and Wilson model including additional information on traveler types and testing whether coefficients differ significantly from the pooled travel cost approach. Loomis et al. (2000) make an improvement to the Parsons and Wilson model by delineating between joint destination and incidental destination visitors and adding in separate slope and intercept shifters for the two sub-groups. Comparing this generalized model to the standard TCM, Loomis et al. find that failing to control for MD visitors increases the CS values by 20%-70% due to overestimating MD travelers' travel costs, highlighting the need to correct for the bias. Although Loomis et al. find the magnitudes of CS vary in accordance with a priori expectations, they do not find statistical significance between the estimated CS values. However, while not statistically different from one another, the authors do note that the differences could still be policy relevant.

Other papers have offered modifications to the original Parsons and Wilson model, most often testing restrictions and attempting to increase explanatory power. Loomis (2006) compares the Parsons and Wilson TCM to a contingent valuation method in valuing the Snake River in Wyoming, and finds similar results between the two methods, both of which are different – though not statistically – from the standard TCM. Hill et al. (2014) compare several specifications of travel cost models for agritourism in Colorado, including Loomis' (2006) approach, and again find results consistent with a priori expectations set by the literature. Each of these studies find similar conclusions but also express a need for more research in overcoming the MD bias. This study attempts to fill this void by addressing the problem by more careful development of the survey instrument rather than simply with econometrics or complex specifications during empirical estimation. The next section lays out a new method for accounting for MD bias by including additional survey questions that elicit the incremental cost of travel from MD visitors to agritourism farms and ranches in the Western U.S.

Methods

This section begins with a review of the economic theory behind the travel cost method and its use to value recreational sites. After detailing the methods used to empirically estimate the travel cost model and the derivation of consumer surplus from the model, I discuss the source of the MD bias and its effect on the CS estimates from the travel cost model. Given the scope of this study, I also discuss multiple sources of heterogeneity in the TCM, and introduce unique survey data to account for both the MD bias as well as some of the heterogeneity in the agritourism industry. Focusing on these different sources of heterogeneity in the agritourism industry will not only lead to more precise CS estimates, but offer insights into how the MD bias might change depending on different trip characteristics and traveler behaviors in the greater TCM literature.

The TCM is based on the microeconomic theory of utility maximization. Equation 3.1 represents an individual's utility maximization problem, where Q is the quantity of agritourism trips taken, $\mathbf{C}$ is the quantity of all other goods and services consumed, $\mathbf{X}$ is a vector of agritourism site characteristics, and $\mathbf{D}$ is a vector of the individual's demographics. The utility function is assumed to be monotonically increasing and is constrained by their income (M), the travel cost associated with the site (p), and the prices of all other goods and services ($\mathbf{r}$).

(3.1) $$Max_{Q,C} \quad U(Q, \mathbf{C}|\mathbf{X}, \mathbf{D}) \quad s.t. \quad pQ + \mathbf{rC} \le M$$

Solving this maximization problem across individuals results in the uncompensated demand for agritourism, shown in equation 3.2, where the quantity of agritourism depends on income, demographics of an individual, the agritourism site characteristics, travel cost, and the prices of all other goods and services.

(3.2) $$Q_{ij} = f(M, \mathbf{D}, \mathbf{X}, p, \mathbf{r})$$

The development of the empirical model makes use of equation 3.3 to estimate an agritourist's demand for the quantity of agritourism sites in the Western U.S. $\mathbf{Q}$ is an nx1 vector of the quantity of agritourism trips taken by n individuals, $\mathbf{M}$ is an nx1 vector of individuals' incomes, $\boldsymbol{\gamma}$ is an nxk_1 matrix where k_1 are demographic variables, $\boldsymbol{\theta}$ is an nxk_2 matrix of site characteristics where k_2 are site descriptors, $\mathbf{p}$ is a vector of travel costs, and $\boldsymbol{\varepsilon}$ is a random error term. As per Rosenthal's (1987) admonition to include substitute prices, I adopt Benson, Watson, Taylor, Cook, & Hollenhorst's (2013) approach using the respondents' meal, shopping, and lodging expenditures from the agritourism site's surrounding area as a proxy for substitute prices.

(3.3) $$\mathbf{Q} = f(\mathbf{M}, \boldsymbol{\gamma}, \boldsymbol{\theta}, \mathbf{p}, \boldsymbol{\varepsilon})$$

Due to the discrete and non-negative nature of the data, estimating truncated Poisson (TP) or truncated negative binomial (TNB) distributions are necessary to achieve consistent and efficient estimates (Grogger & Carson, 1991; Martínez-Espiñeira & Amoako-Tuffour, 2008). While the Poisson is commonly used in many count data models, it can sometimes lead to an unrealistic statistical assumption termed equi-dispersion where the sample mean is equal to the sample variance. While the equi-dispersion does not affect the consistency of parameter estimates, assuming a Poisson distribution when the data is over-dispersed can lead to the standard errors being deflated, thus increasing the probability of committing a type one error in significance testing (Martínez-Espiñeira & Amoako-Tuffour). After conducting a likelihood ratio test and finding significant over-dispersion ($p < 0.01$), I adopted the TNB distribution to ensure efficiency within parameter estimates (Grogger & Carson, 1991).

Consumer surplus is defined geometrically as the area under the demand curve but above the market price, TC_0. This can be calculated by taking the definite integral of the demand curve, defined between the choke price and the market price. In the travel cost model, the travel cost represents the market price of the activity, while the choke price is assumed to be asymptotic to the price axis and approaches infinity as the quantity of trips falls to zero.

Due to the close relationship between the Poisson and the negative binomial distributions, the first moment of the negative binomial model is simply λ. For estimation purposes, an exponential parameterization ($\exp(\beta x_i)$) is frequently used to restrict the dependent variable, in this case quantity of trips, to positive values. For these reasons, in order to derive CS, the integral over the demand curve can be written as: $CS = \int_{P_0}^{P_1} \exp(\beta x_i)\, dp$, since $E[Q_i|X_i] = \lambda_i = \exp(\beta x_i)$. Given this set up, the per person CS for the travel cost model can be derived as such:

(3.4)
$$CS = \int_{TC_0}^{\infty} \exp(\beta_0 + \beta_1 TC)\, dp$$

$$CS = \frac{\exp(\beta_0 + \beta_1 TC)}{\beta_1}\Big|_{TC_0}^{\infty}$$

$$CS = 0 - \frac{\exp(\beta_0 + \beta_1 TC_0)}{\beta_1}$$

$$CS = -\frac{\widehat{\lambda_i}}{\beta_1}$$

Dividing equation 3.4 the expected number of trips taken by individual *i* ($\widehat{\lambda_i}$) results in the average per trip per person consumer surplus:

(3.5)
$$CS = -\frac{1}{\beta_1}$$

Finally, multiplying equation 3.5 by the total number of recreational users gives the total CS for the recreational site.

Traditionally, travel cost measurements are either extracted from secondary data sources where the respondent's home and destination zip codes are known, or it is elicited from survey respondents using a similar beginning- and end-point method. The issue with these measurements is that they do not account for the differences in travel behavior between respondents. While a start-to-finish travel measurement might be justifiable for primary purpose travelers who treat the site in question as the primary purpose of their travel, this measurement is incorrect for multi-destination travelers who may view their visit to the site in question as a secondary or incidental stop. Referencing figure 3.1, if a traveler's primary destination is not the recreational site, but instead some other destination at *D*, then when considering their travel cost to the recreation area, the first leg of their trip, between *A* and *B*, is a sunk cost. In fact, the only relevant travel cost for these multi-destination visitors is the incremental travel cost from their primary path of travel at *B* to the site in question at *C*. Empirically, this can be thought of as a measurement error bias where the travel cost is inflated by some sunk cost (*A* to *B*) for multi-destination travelers.

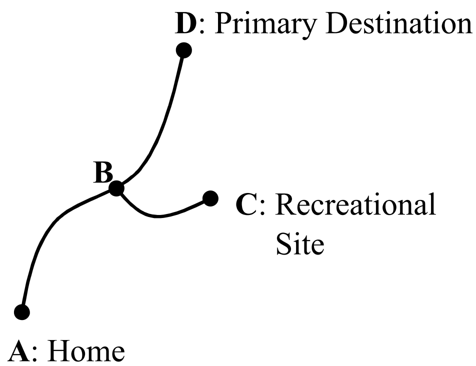


Figure 3.1: Measuring travel costs for multi-destination visitors

The measurement error bias from incorrectly measuring the MD visitors' travel cost is represented in equation 3.6 where the travel cost vector ($\widetilde{\boldsymbol{TC}}$) is made up of the true travel cost vector ($\boldsymbol{TC}$) plus some positive measurement error ($\boldsymbol{u}$) from including the respondent's sunk costs of travel.

$$\widetilde{TC}_i = TC_i + u_i \tag{3.6}$$

In a linear regression model using ordinary least squares, this measurement error can easily be shown to inflict an attenuation bias in the relevant parameter estimates through a violation of the independence between the explanatory variables and a composite error term. However, in the present analysis, there is no closed form for the parameter estimator, and instead the Newton-Raphson search algorithm is used to find the parameter values that maximize the log-likelihood equation for the TNB. Regardless, a heuristics argument that starts with the simplest specification of the travel cost model with just one independent variable (TC_i) and parameter β_1, that the parameter estimate will be biased downward due to overmeasurement in

TC_i. That is, for a specific quantity of trips, $\overline{\lambda_i}$, in order for the equality to hold, the travel cost parameter, $\widetilde{\beta_1}$, must take on smaller values relative to β_1 to account for the inflated TC variable from the inclusion of sunk costs. In addition, Cameron & Trivedi (2013) point out that if the variance of an independent variable increases due to measurement error without any change in the range of the dependent variable, then the independent variable must produce a smaller effect.

$$\overline{\lambda_i} = \exp(\widetilde{\beta_1}\widetilde{TC_i}) \qquad \overline{\lambda_i} = \exp(\beta_1 TC_i)$$

$$\ln(\overline{\lambda_i}) = \widetilde{\beta_1}\widetilde{TC_i} \qquad \ln(\overline{\lambda_i}) = \beta_1 TC_i$$

$$\widetilde{\beta_1} < \beta_1$$

Comparing this result to the CS calculation shown in equation 3.5, it is not difficult to see that the attenuation bias in the travel cost parameter leads to overestimating the CS. This overestimation due to the inclusion of sunk costs in the travel cost measurement will be tested in the results section using Wald tests between individual pairs and groups of TC parameters between the incrementally and cumulatively measured TC parameters. If the tests are statistically significant, this will be evidence of statistically significant measurement error bias from including sunk costs of travel.

While the bias in the parameter estimates are due to the inclusion of sunk costs, there are a few empirical sources that may lead to different magnitudes of bias. The first is the share of multi-destination visitors in the sample. If the share of multi-destination visitors to a particular recreational site is relatively small, then the CS estimates may not suffer greatly from the attenuation bias. Second, if the sunk costs of travel are relatively small compared to the incremental travel cost, then, again, the difference between the parameter estimates may not be significant. A scenario where this might occur is in estimating recreational sites close to urban areas, where many of the visitors may be from nearby, meaning even if they are a multi-destination visitor, their sunk costs of travel will be relatively small. The last empirical source of the MD bias deals with the relative price of the recreational activity. The more expensive the ticket price is for a recreational activity, the less passersby will be willing to travel out of their way to visit the site. This relative difference between the sunk cost of travel and the relatively small incremental travel cost may lead to larger magnitudes of bias in the travel cost parameter estimates.

Accounting for Industry Heterogeneity

In addition to the exploration of the multi-destination bias, this analysis develops a somewhat novel approach to incorporate the heterogeneity of agritourism sites across space into the empirical model. This was done to identify comparative advantages in the agritourism industry, as well as to explore how the MD bias differs depending on the travel experience. In the past, the TCM has mainly been applied to a specific site or a small set of sites, so attempting to use the TCM to value an entire regional industry presents additional complexities. To account for differences in agritourism demand across the Western U.S., three sources of heterogeneity are explored: agritourism activities, regional, and the trip types as discussed above (primary purpose vs. multi-destination).

The four agritourism activities included in the empirical model are direct to consumer sales, entertainment and events, outdoor recreation, and educational activities. These activity groups were identified by conferring with an advisory board of agritourism industry stakeholders. While the value consumers place on specific activities within these generalized categories likely vary, consumers are assumed to view activities within these categories as similar experiences. Summary results and a greater discussion of these activities are presented in the next subsection outlining the survey instrument and data.

TCM studies using secondary data use geographical zones as a means to calculate the number of travelers visiting a site by multiplying the zone's population by the estimated share of visitors for the recreation site in question. The motivation to define agritourism zones is partly out of practicality, and partly due to the nature of the research question. From the practicality standpoint, there need to be few enough zones to ensure there are enough degrees of freedom to achieve statistical significance in the multivariate regression analysis. Effectively, zones allow the researcher to reduce the number of agritourism destinations to geographical units that capture similar flavors of agritourism. In addition, since this chapter attempts to explore differences in consumer surplus across space and agritourism activities, defining agritourism zones, as opposed to other geographical units, allows the researcher to capture the heterogeneity of different agritourism climates across regions.

In order to address the heterogeneity of agritourism experiences across space, I conducted a principal components analysis (PCA) on a set of county level explanatory variables related to the agritourism activity in a county. Modeling the covariance structure between the explanatory variables with some latent factor, allowed me to develop an "Agritourism Score" for each county in the Western region. PCA is commonly used to reduce the dimensionality of a data set or to describe some latent factor. Eigenvectors essentially measure the direction of each dimension of the data set, and each eigenvector has one eigenvalue. Eigenvalues are measures of the variance within an eigenvector, or a specific dimension of the data, and since they can be made to be orthogonal to one another, they can be used to explain dimensions of the data that are independent of each other.

Thirteen county-level variables were selected that were hypothesized to explain the agritourism climate for each county. These variables were selected based on results from chapter two and from findings in the literature (Van Sandt et al, 2018; Hill et al, 2014). (Chapter two is an excellent source to draw from considering its objectives were to explore factors relating to the propensity to adopt an agritourism establishment as well as agritourism site and locational characteristics that create a greater draw for tourists to participate in the recreational experience.) The PCA variables are presented in table 3.2 along with their factor loadings. Factor 1 had the highest eigenvalue (1.2769)[6] and the factor loadings fit a priori expectations in terms of their relationship to the latent variable which was designed to simply be a measure of a county's agritourism climate. It is important to note here that because the PCA describes a latent variable, it would be remiss to assume a specific definition for the underlying latent variable. Moving forward, the score which presumably captures the covariance structure of the latent variable will be referred to as the "Agritourism Score" of a county. The factor loadings indicate how much of the variation in that factor loading is attributable to that specific variable. Note these are simply relationships among the data and imply nothing about the causality of each variable with the latent factor.

Table 3.2 Principal Components Analysis Factor loadings

Variable	Factor Loading
Average Farm Sales	-0.1069
Population (2010)	0.2725
Agritourism Revenue/Farm or Ranch	0.0213
Total Agritourism Revenue	-0.1826
Number of Agritourism Farms	0.2922
Time to a Population of 250k	-0.5788
Natural Amenities Scale	0.2699
Farm Dependent (2003)	-0.2236
Recreation Dependent (2004)	0.0881
Scenic Byways/hundred sqm	0.5181
Interstates/hundred sqm	0.0943
Time to National Parks Service Asset	-0.4744
Agritourism Hot Spot Residuals	0.3226

Factor 1 had an eigenvalue of 1.2769

[6] Two factors from the principal components analysis had eigenvalues above one.

The factor loadings in table 3.2, were used to create an indexed value or "Agritourism Score" for each county. Given this score, a choropleth map was developed across counties in the Western U.S., shown in figure 3.2. While it is clear that there are many areas that could be seen as unique agritourism clusters, part of the motivation to create agritourism zones is to reduce the spatial dimensionality and save degrees of freedom in the travel cost model. After comparing like-agritourism scores in figure 3.2, seven agritourism zones were defined: Northwest, Northern Plains, Central California, Central, Mountain, Southwest, and Southern Plains.

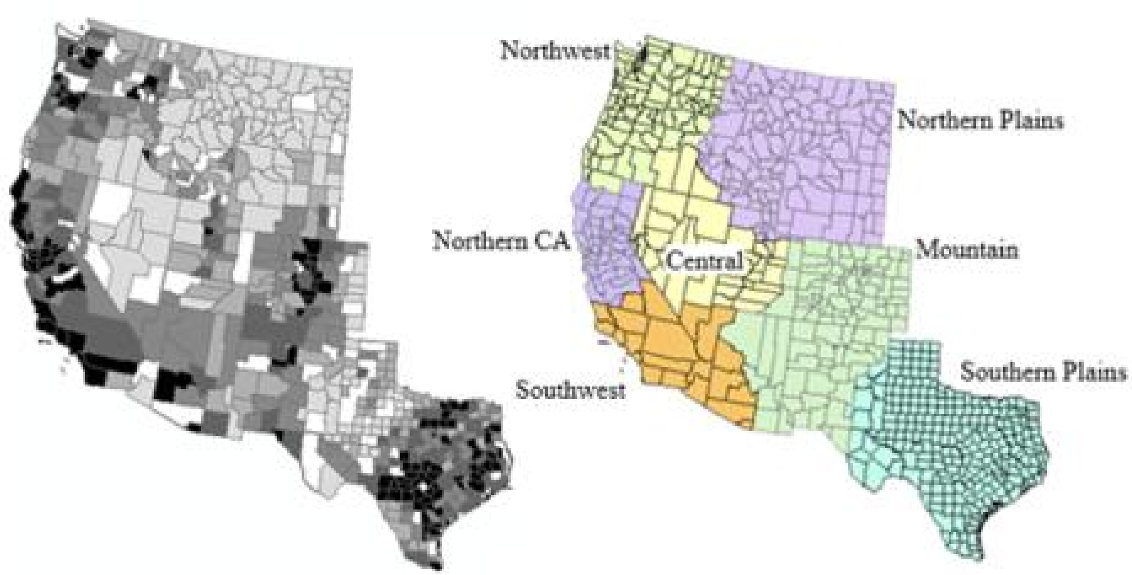


Figure 3.2 Principal Component Analysis Scores and Agritourism Zones
Darkly shaded counties in the PCA map can be interpreted as being more conducive to agritourism.

While the method of using the PCA to motivate how the TCM zones should be defined is somewhat ad hoc, it was perceived as a preferred method considering the alternatives. In economics it is common to use counties, states, or regions as geographical units of analysis, but these political boundaries are not appropriate in the present context. These geographical units may make sense for understanding regional economic growth or observing spatial spillovers between local economies, but there is nothing to indicate that different flavors of agritourism should follow the political borders drawn on maps. Previous research has shown that agritourism is likely dependent on climate, natural amenities, and populations, all of which are more or less continuums across space and do not follow any particular unit of geography that is commonly used.

Empirical Model

In order to observe the consequences of restrictive travel cost model specifications, I developed four model specifications, each less restrictive than the one before. The first model represents a standard TCM with no delineation between traveler type, region, or agritourism activities. Model two accounts for greater regional and agritourism activity heterogeneity by interacting TCs with regions and activities, but omits traveler types. Model three is the opposite of model two, and closely resembles the Parsons and Wilson method by incorporating traveler heterogeneity, but not regional or agritourism activity types. The final (fourth) model offers the greatest flexibility in terms of intercept shifts and slope rotations across the three sources of heterogeneity described above: agritourism activity, region, and traveler type. Likelihood ratio tests for model fit will be used to test the explanatory power of each specification. A summary of these eight models is presented in table 3.3 in the results section.

The empirical specification employed for model four is presented in equation 3.6, and while not explicitly shown, the reader can deduce the specifications of models one through three from this equation and table 3.3. In model four, equation 3.6, traveler types were interacted with both regions and activities, however, perfect multicollinearity is avoided since activities offered at the site are not mutually exclusive. Note the constant is suppressed making the reference group PP travelers participating in some activity or set of activities; a choice made due to a relatively flat likelihood function that caused issues with convergence in the iterative maximum likelihood estimation process. By suppressing the constant, the heterogeneity in the data is effectively forced to be captured by the variable coefficients rather than being absorbed by the constant.

The fully flexible form utilized in model four can be written as:

$$\begin{aligned} \boldsymbol{Q(trips)} = {} & \boldsymbol{\beta_1 (TT * PP) + \beta_2 (TT * MD) + \beta_3 Secondary + \beta_4 Incidental + \beta_5 Inc} \\ & \boldsymbol{+ \beta_{k1} Activities + \beta_{k2} Activities * (PP * TC) + \beta_{k3} Activities} \\ & \boldsymbol{* (MD * TC) + \beta_{k4} Regions * (PP * TC) + \beta_{k5} Regions * (MD * TC)} \\ & \boldsymbol{+ \beta_{k6} \overline{D} + \varepsilon} \end{aligned} \tag{3.6}$$

This fully flexible form in equation 3.6 allows for different demand curve intercepts across traveler types (β_3 and β_4) and activities (β_{k1}), as well as rotations in the demand curve for each activity (β_{k2} and β_{k3}) and region (β_{k4} and β_{k5}) across traveler types. As before, $\boldsymbol{\overline{D}}$ represents trip and traveler specific characteristics such as surrounding area expenditures, age, whether or not the respondent engaged in multiple activities, and the natural amenities index for the county the agritourist visited. Each of the preceding three models are special cases of equation 3.6, and although the constant is suppressed, equation 3.6 theoretically collapses into the other three models if none of the coefficients within k_4 and k_5, and k_2 and k_3 are significantly different from one another.

Survey Data

The data for this research was compiled using a survey distributed by Taylor Nelson Sofres (TNS) in late April, 2015, using their pre-recruited panels, to a stratified regional and national sample of travelers, with a focus on agritourists in the Western region. The Western region was defined as all states west of, and including, Montana, Wyoming, Colorado, and Texas, but excluding the noncontiguous states of Alaska and Hawaii.

The survey varied in length depending on whether the respondent was randomly selected to receive one of four versions. The two survey versions used in this essay were: 1) most recent trip involving agritourism and 2) longest trip involving agritourism, totaling 1,001 observations. Details on the specific questions in the two versions will be discussed in greater detail below, but regardless of the survey version, the primary purpose of survey was to explore the travel behavior of agritourists, determine how agritourists differ from other travelers, and test a new method for correcting MD bias. In general, questions in the survey asked for distances/time traveled, dollars spent, the primary purpose of the trip, agritourism activities participated in, and other agritourism or general travel-related questions.

The benefits of using a research company and its standing consumer panel are twofold; a high response rate is achieved due to the company's pre-recruitment and incentives, and it is more likely to achieve a balanced sample due to TNS' large and established panel. The initial goal was to survey 1,600 travelers who specifically visited agritourism sites in the Western region (Texas, Colorado, Wyoming, Montana, and all states westward), however, TNS' initial screening showed too low an incidence of the participation in the specific definition of Western agritourism[7] among travelers in the panel (around 8% instead of the expected 10%) to meet the sample goal. Given this lower than expected incidence, TNS had to both extend their potential sample by using their panel partners to achieve the desired quota of agritourists in the West, and make part of the sample based on national travelers who did not visit agritourism sites, hence there were two additional versions of the survey created that are not used or discussed in the research presented in this essay.[8]

[7] "Agritourism is any paid for educational or recreational activity that takes place on an operating farm nursery, or some other agricultural production site. Examples of agritourism are wine tastings, a corn maze or dude ranch."

[8] Due to the fact that TNS had a goal a specific number of observations, the respondents of the non-agritourism survey versions cannot be used in a two-stage probit as is commonly done to control for participation bias (e.g. Hill et al. 2014). This instead led to the use of a truncated model to achieve non-bias parameter estimates.

The two main versions of the survey targeted the respondent's most recent trip or longest trip in the past year, since both types of trips are of interest in understanding key traveler behavior for that sector. For example, the longest trip a family goes on frequently encompasses more people, planning, and activities, and may be important given the distances one needs to travel to see some Western sites, but the "most recent" may provide a better snapshot of current or frequent agritourism travel patterns. The "longest" and "most recent" versions of the surveys were randomly split 50-50 within the self-designated agritourist group to assure a balanced sample.

While some variables such as demographics and type of agritourism site were taken directly from the survey results, other variables had to be constructed due to obstacles encountered during the data collection stage. Most travel cost studies revolve around one or several specific sites allowing them to use an individual or zonal travel cost approach depending on if they have primary or secondary data (Loomis & Walsh, 1997). However, given the task of valuing an entire industry, directly eliciting the respondent's quantity of trips to the particular site they answered the survey questions on was not feasible. In some cases, it is probably unlikely that they visited the site in question more than one time. Therefore, for this chapter that focuses on the size of a tourism sector comprised of relatively similar sites, the final quantity of trips variable in the empirical model represents the number of trips an individual took to an agritourism site *like* the one they answered the survey questions on (direct sales, event or entertainment, outdoor recreation, education, or other), while on a trip (day or overnight) *like* the one they answered the survey questions on. Several other variables such as total number of trips and hours traveled were cross-checked to ensure the number of trips were consistent with the rest of the respondent's answers. Due to this evaluation of "like" agritourism trips using individual travel data, this approach may be referred to as a hybrid approach; that is, this approach draws from both the individual and zonal TCMs. The summary statistics of the relevant variables are presented in table 3.2.

While the previous literature, such as Loomis et al. (2000) and Parsons & Wilson (1997), have included dummy variables and dummy interaction terms to capture shifts and rotations for the demand curve, the incremental mileage traveled off of the traveler's path to their primary destination has never been directly requested from travelers. Inquiring whether the agritourism visit was a primary, joint, or incidental visit, as well as the additional miles traveled to the agritourism site from their primary path of travel, enables the researcher to disentangle the true travel cost incurred by the traveler to visit the site. To assure the respondents understood the distance requested, a graphic was included in the survey visually depicting the intention behind the incremental travel cost question for MD travelers (much like figure 3.1).

Due to the desire to include a comprehensive set of agritourists, respondents to this survey were not thrown out of the sample if they did not travel by vehicle as was done in Carpio et al. (2008). Instead each respondent was asked what mode of transportation they traveled by, allowing for more precise travel costs calculations for each respondent depending on their mode of transportation. While different modes of transportation may vary in more ways than just cost, including travel time as an independent variable captures another significant characteristic of different travel modes. In a perfect world, each mode of transportation would be included in the travel cost model as a dummy variable, or even an interaction variable with travel costs, but based on limited degrees of freedom for some modes, these were not entered in as explanatory variables in this model. An investigation into how travel mode affects consumers' travel behavior is left to future research.

The majority of the sample traveled via their own vehicles (620 obs.) and rental vehicles (123 obs.). Travel costs for these modes of transportation were calculated using traditional methods in the literature: namely, multiplying the round-trip miles by $0.14/mile. Loomis and Walsh (1997) suggest using $0.12/mile, however, after crosschecking mileage costs with the American Automobile Association (AAA, 2014), $0.14/mile was observed to be a more current per-mile cost for automobiles. In addition to the per-mile travel, the average rental fee for the state each respondent visited on their agritourism trip was calculated assuming each respondent chose an economy class three-day rental.

The next most popular mode of transportation was by airplane (65 obs.). To calculate this travel cost, each respondent's home state and destination state was first observed. Using this information, the average airfare per air mile for each of the hubs, and then each of the states, to come up with an estimate of the traveler's expected airfare (U.S. Dept. of Transportation, 2014). Although this is a very general measure of airline travel costs, it allows the researcher to overcome difficulties from making strict assumptions about how far in advance the traveler bought the airline ticket, what season the traveler flew in, and what airport the traveler chose. However, some traveler heterogeneity is lost due to these generalities. Finally, any multi-destination miles traveled were assumed to have marginal cost of $0.14/mile to remain consistent with the other automobile travel cost rates.

Motorcycles (7 obs.), tour busses (27 obs.), and trains (9 obs.) had smaller frequencies than automobiles and planes as reported modes of transportation. The travel costs for motorcycles were calculated the same way as for private vehicles except that $0.08/mile was used instead of $0.14/mile to remain consistent with AAA's gas mileage estimates. In order to calculate the travel costs for respondents who traveled by tour busses, data was collected from several tour bus companies in the West and estimated their per day costs. Drivers were assumed to drive an average of 6 hours/day (the legal maximum is 10 hours/day), resulting in the per hour cost, which was then multiplied by the respondent's travel time. This calculated per hour travel cost was compared with some day trip tours that explicitly gave the number of hours traveled in a trip, and the two measures were comparable. This method may not be accurate for incidental visitors who traveled by bus, however, the sample only contained one observation for a multi-destination agritourist on a tour bus. Finally, the travel costs for trains was calculated by observing Amtrak prices for 15 major train routes in the West, varying in distance from 88.7 miles to 649 miles. Per mile travel costs were simply calculated using an average of these commuter and cross-country train route fares/mile. Any multi-destination miles traveled were again assumed a constant travel cost rate of $0.14/mile.

As mentioned previously, travel time plays an important role in travelers selecting their mode of travel, but it also plays an important role in accounting for the opportunity costs of travel (see Bockstael et al, 1987). In order to calculate the opportunity cost of travel for each respondent, the researcher multiplies the roundtrip travel time in hours by a fraction of the wage rate. There is no consensus in the literature as to what fraction of the wage rate to use, but the vast majority of the TCM literature uses either one-fourth or one-third of the wage rate. After experimenting with both fractions of the wage rate, the larger fraction of the wage rate was used due to negligible differences in the model's results. Since the income variable is categorical, the midpoint of each income group was used to calculate the wage rate, assuming a 40-hour work week and no time off.

About 27% of the sample were multi-destination visitors who either treated the agritourism site as a secondary destination (19%) or an incidental destination (8%). To save

Table 3.2 Summary Statistics Observations: 852				
Variable	Mean	Standard Deviation	Min	Max
Agritourism Trips	2.5716	4.5049	1	80
PP Travel Cost	24.3011	54.5607	0.08	769.5876
MD Travel Cost	1.4875	7.4742	0	93.33333
Travel Time (minutes)	106.8179	76.0367	0.01	1251.282
Primary Destination	0.7336	0.4424	0	1
Secondary Destination	0.1913	0.3936	0	1
Incidental Destination	0.0751	0.2637	0	1
Female	0.5704	0.4953	0	1
*Income**	4.1279	2.0702	1	8
Age	37.3416	14.0050	18	84
Agritourism Zones (figure 3.2)	Percent of Agritourism Sites Visited			
Mountain	21.13%			
Southwest	26.17%			
Central	1.64%			
Northern Plains	7.87%			
Northern California	19.25%			
Northwest	12.32%			
Greater Texas	11.62%			
Agritourism Activities	Percent of Agritourism Sites Visited			
Direct Sales	33.80%			
Entertainment and Events	43.43%			
Outdoor Recreation	43.19%			
Education	32.98%			
**Income has 8 categories: 1=under $30k, 2=$30k-$40k, 3=$40k-$50k, 4=$50k-$75k, 5=$75k-$100k, 6=$100k-$125k, 7=$125k-$150k, 8=$150k and over*				

degrees of freedom, both secondary and incidental visitors were classified as MD visitors. The potential for a large positive bias in travel costs is immediately apparent considering the disparity in mean incremental travel costs between PP and MD visitors ($24.30 vs. $1.49). *Entertainment and Events* and Outdoor Recreation were slightly more popular than *Direct Sales* and *Education* activities, but the similar share of respondents who participated in each activity category implies that the categories do a relatively decent job at capturing the industry's activity heterogeneity. The *Central* and *Northern Plains* regions were aggregated into the *Central/Northern Plains* region due to their low visitation rates, which are likely due to their relative rurality and lack large population centers where potential agritourists may set out from. The *Southwest* had the highest visitation level, followed by the *Mountain* and *Northern California* regions.

Results

All models were estimated using a constant-suppressed TNB via maximum likelihood estimation. While heteroscedasticity was not tested for, each model made use of White's robust standard errors to improve efficiency without affecting the consistency of parameter estimates. Each model was tested for multicollinearity using an un-centered variance inflation factor (VIF). Significant levels of multicollinearity were detected in model eight for two MD travel cost interaction variables,[9] but due to general significance across the variables of interest and the fact that multicollinearity does not affect the consistency of parameter estimates, the presence of multicollinearity between these two interaction terms should not affect the interpretation of results. Future studies attempting to value heterogeneous industries may be able to circumvent this issue by collecting larger samples. The average VIF for the model was 2.96, well below conventional thresholds for multicollinearity.

Table 3.3 Travel Cost Models of varying degrees of flexibility

Variables	Model 1 Standard TCM	Model 2 TC Interactions	Model 3 Parsons & Wilson	Model 4 Fully Flexible Form
Travel Cost	**X**	**X**	**X**	**X**
TC*Activities		**X**		**X**
TC*Regions		**X**		**X**
TC*MD			**X**	**X**
Flexibility	None	Experience	Traveler	Full
Wald Chi^2	90.36***	113.38***	95.69***	149.76***
Pseudo LL	-1,353.94	-1,337.68	-1,353.00	-1,329.03

Each model was estimated using incremental TC calculations.
*** Model is significant at the 1% level.

---

[9] VIFs for model eight: $MDTC * Mountain$ VIF= 13.38, $MDTC * Ent.\&Events$ VIF=8.79

Table 3.3 shows the four model specifications with varying degrees of flexibility in their treatment of activities, regions, and travelers. Using likelihood ratio tests for model fit, models two, three, and four, were all statistically significant improvements on model one at the one percent level. Based on the likelihood ratio tests, model four also offers a better fit than model three at the one percent level, but is not a significant improvement on model two. This indicates that including regional and activity heterogeneity accounts for a greater degree of variation in the dependent variable than solely accounting for traveler differences. The Wald Chi^2 statistics and pseudo loglikelihood values shown in table 3.3 also support the conclusions of the tests. However, it should be noted that the travel cost interaction variables' coefficients on two activities (Events and Entertainment ($P = 0.0108$) and Outdoor Recreation ($P = 0.0134$)) and four regions (Southwest ($P = 0.0480$), Mountain ($P = 0.0759$), Central/Northern Plains ($P = 0.0548$), and Southern Plains ($P = 0.0086$)) were significantly different between PP and MD travelers.

The effect of the measurement error can be observed through three different lenses: 1) the effect on the estimated travel cost coefficients, 2) the effect on part worth CS estimates, and 3) the effect on the overall valuation of agritourism in the Western US. Each of these will be discussed in turn before discussing the applied implications of the varying CS values across regions, travelers, and activities.

Wald tests were conducted on individual travel cost interaction terms, as well as on groups of interaction terms, between model four estimated with and without measurement error in the travel costs. As a group, the regional travel cost interaction parameters for MD travelers were significantly different when measured with and without measurement error ($P = 0.0225$), however the PP regional travel cost interaction parameters did not differ significantly. Similarly,

Table 3.4 Travel Cost Model 4 – Constant Suppressed Truncated Negative Binomial Observations: 852		
Variable	Coefficient	Robust Std. Errors
Travel Time - Prim	0.0015**	0.0006
Travel Time - MD	0.0054	0.0056
Secondary	-0.3920*	0.2330
Incidental	-0.2670	0.2750
Female	-0.2780	0.1850
Ln(Income)	-0.5550***	0.1570
Age	0.0123*	0.0072
Meals Expenditures	0.0006*	0.0003
Rec. Expenditures	-0.0004	0.0007
Shopping Expenditures	0.0001	0.0004
Agritourism Expenditures	-0.0006*	0.0003
Natural Amenities	-0.0458*	0.0256
Activity Types		
Direct Sales	-0.0122	0.1870
Entertainment/Event	-0.6820***	0.2420
Outdoor recreation	0.3800	0.2450
Education	-0.1820	0.2030
Multiple Activities	0.9160***	0.2900
Prim TC * Direct Sales	0.0026	0.0023
Prim TC * Ent./Events	0.0092***	0.0031
Prim TC * Outdoor Rec.	-0.0030	0.0023
Prim TC * Education	0.0030	0.0030
MD TC * Direct Sales	0.0524*	0.0307
MD TC * Ent./Events	0.0698***	0.0242
MD TC * Outdoor Rec.	0.0484**	0.0207
MD TC * Education	-0.0013	0.0334
Regional Zones		
Prim TC * Mountain	-0.0041	0.0034
Prim TC * Southwest	-0.0045*	0.0025
Prim TC * Central & NE	0.0057	0.0080
Prim TC * Northern CA	-0.0100**	0.0044
Prim TC * Northwest	-0.0054	0.0073
Prim TC * Texas	-0.0056	0.0036
MD TC * Mountain	-0.0736*	0.0396
MD TC * Southwest	-0.1000**	0.0488
MD TC * Central & NE	-0.1870*	0.0994
MD TC * Northern CA	0.0430	0.0832
MD TC * Northwest	-0.0662	0.0482
MD TC * Texas	-0.1200***	0.0441
Alpha – over-dispersion parameter: 1.43***		
Wald Chi^2: 149.76*** * Significant at 1% level ** Significant at 5% level *** Significant at 10% level		

the activity travel cost interaction parameters for MD travelers were significantly different ($P = 0.0229$), but the corresponding interaction terms' parameters for PP travelers were not significantly different. This is to be expected considering the measurement error from including the sunk costs of travel only pertains to the MD travelers. Taking a closer look at these regional and activity interaction terms for the MD travelers shows varying effects of the bias depending on the region and activity. Still using the Wald test, there were significant differences in the parameters for Events and Entertainment ($P = 0.0189$) and Outdoor Recreation ($P = 0.0284$) after accounting for measurement error, as well as between the Southwest ($P = 0.0182$), Mountain ($P = 0.0734$), and Southern Plains ($P = 0.0873$) regions. This indicates that the sunk costs of travel, and therefore the consumer's travel behavior, differs significantly across regions and activities.

Tables 3.5 and 3.6 show the part-worth CS estimates for each activity and region for PP and MD travelers respectively. Note these are part-worth CS values since each agritourism experience may include more than one activity. The standard errors for the CS estimates were calculated using a procedure similar to that used in Adamowicz, Fletcher, & Graham-Tomasi (1989). This procedure involves using Monte Carlo simulations with 500 iterations to develop confidence intervals around the part-worth CS estimates presented in tables 3.5 and 3.6. All PP partial CS estimates and MD partial CS estimates, except for MD travelers participating in Events and Entertainment in the Mountain region, are significantly different from zero at the five percent level. Superscripts denote significance (at the five percent level) between CS estimates for other CS estimates in the same column and/or row to facilitate comparisons. Nearly all PP and MD part-worth CS estimates that were estimated using the incremental travel costs were significantly different from those estimated using the cumulative travel costs except for six part-worth CS estimates.[10]

The fact these CS estimates are part-worth values to some extent explains why some of the part-worth consumer surplus estimates are negative. Some negative partial consumer surplus estimates may reflect a contrast to another region's strong comparative advantage in a specific activity. For example, Northern California has the only positive PP consumer surplus value for entertainment and events most likely due to the region's notable visibility and reputation for winery tastings, events, and tours. Moreover, Northern California's events and entertainment generate the highest consumer surplus in the entire model indicating, given the choice, PP travelers seem to see all other wine regions as inferior destinations. The negative MD partial consumer surplus values for Northern California add to this narrative showing the commitment and planning required to visit the relatively more expensive agritourism activities in the nation's most renowned wine region. In contrast, it seems agritourism sites in the Central/Northern Plains region are more likely to attract MD visitors.

Tables 3.7 and 3.8 show the relative bias generated by incorrectly measuring the MD travel costs in model four. The percent bias was calculated as the percent change between the CS estimates measured with and without the sunk costs of travel, and can be interpreted as a relative degree of bias in each partial consumer surplus estimate as a result of incorrectly measuring the travel costs. As expected from the Wald tests described above, the PP part-worth consumer surplus estimates are somewhat biased but the differences are not nearly as large as the biases in

[10] Partial consumer's surplus estimates for PP travelers that were not significantly different across models seven and eight were: Direct to Consumer Sales (DTC) and Outdoor Recreation (Outdoor Rec.) for the Central/Northern Plains region; DTC, Entertainment and Events (Ent. & Events), and Education (Edu.) for the Northwest region. MD travelers' partial consumer's surplus for Ent. & Events in the Northwest region was not significantly different across models seven and eight.

Table 3.5 Primary Purpose Visitors' Per Person Partial Consumer Surplus Estimates						
Activities	Mntn[1]	SW[2]	CtNP[3]	NoCal[4]	NW[5]	SoPlns[6]
DTC[A]	$ 657.89[ALL]	$ 518.13[ALL]	$ - 121.21	$ 135.69[ALL]	$ 362.32[ALL]	$ 338.98[ALL]
Std. Error	*4.980704*	*2.194955*	*1.052026*	*0.308093*	*6.091764*	*0.900318*
Ent. & Events[B]	$ - 196.85	$ - 214.13	$ - 67.34	$ 1,298.7[ALL]	$ - 260.42	$ - 273.97
Std. Error	*0.61573*	*0.529253*	*0.322723*	*9.276213*	*3.678631*	*1.928271*
Outdoor Rec.[C]	$ 140.65[ALL]	$ 132.98[ALL]	$ - 375.94	$ 77.16[ALL]	$ 119.76[ALL]	$ 117.10[ALL]
Std. Error	*0.227238*	*0.131972*	*12.73976*	*0.10614*	*0.650379*	*0.142181*
Education[D]	$ 917.43[ALL]	$ 666.67[ALL]	$ - 115.21	$ 144.09[ALL]	$ 429.18[ALL]	$ 396.83[ALL]
Std. Error	*10.75081*	*4.752308*	*0.941027*	*0.373352*	*8.699483*	*1.378283*
Superscripts indicate the significance of the CS estimate from other activities (A,B,C,D) and other regions (1,2,3,4,5,6)						

Table 3.6 Multi-destination Visitors' Per Person Partial Consumer Surplus Estimates						
Activities	Mntn[1]	SW[2]	CtNP[3]	NoCal[4]	NW[5]	SoPlns[6]
DTC[A]	$ 47.17[236,C]	$ 21.01[1356,BD]	$ 7.43[1256,D]	$ - 10.48	$ 72.46[236,D]	$ 14.79[1235,BD]
Std. Error	*5.905498*	*0.40997*	*1.197091*	*4.906431*	*0.789432*	*0.376891*
Ent. & Events[B]	$ 263.16	$ 33.11[36,ACD]	$ 8.53[26D]	$ - 8.87	$ - 277.78	$ 19.92[23,ACD]
Std. Error	*2.383188*	*0.338322*	*1.863261*	*25.1675*	*0.538602*	*0.502403*
Outdoor Rec.[C]	$ 39.68[236,D]	$ 19.38[1356,BD]	$ 7.22[1256,D]	$ - 10.94	$ 56.18[236,D]	$ 13.97[1235,BD]
Std. Error	*44.15173*	*1.041223*	*0.585101*	*0.967293*	*3.839151*	*0.232386*
Education[D]	$ 13.35[236,AC]	$ 9.87[ALL]	$ 5.31[ALL]	$ - 23.98	$ 14.82[236,AC]	$ 8.24[ALL]
Std. Error	*7.331792*	*1.586491*	*0.468797*	*0.678145*	*10.9298*	*0.197574*
Superscripts indicate the significance of the CS estimate from other activities (A,B,C,D) and other regions (1,2,3,4,5,6)						

Table 3.7 Relative Bias in Primary Purpose Per Person Consumer Surplus Estimates

Activities	Mountain	Southwest	Central & N. Plains	Northern California	Northwest	Southern Plains	Average
DTC	-10%	-6%	3%	-12%	-6%	-26%	-10%
Ent. Events	1%	2%	3%	-51%	1%	-28%	-12%
Outdoor rec	-9%	-8%	-10%	-11%	-8%	-15%	-10%
Edu.	-10%	-4%	4%	-12%	-6%	-28%	-9%
Average	-7%	-4%	0%	-21%	-5%	-24%	-10%

Table 3.8 Relative Bias in Multi-Destination Per Person Consumer Surplus Estimates

Activities	Mountain	Southwest	Central & N. Plains	Northern California	Northwest	Southern Plains	Average
DTC	-403%	-278%	538%	1,355%	-177%	96%	188%
Ent. & Events	-134%	-198%	593%	3,418%	84%	66%	638%
Outdoor rec	881%	-401%	352%	633%	-313%	63%	202%
Edu.	1,128%	-844%	451%	302%	-1497%	155%	-51%
Average	368%	-430%	484%	1,427%	-476%	95%	244%

the MD partial consumer surplus estimates. On average, the PP part-worth consumer surplus estimates are depressed by about 10% whereas the MD partial consumer surplus estimates are inflated by 244%. The latter result should be expected since on a given trip the incremental TC necessarily has to be lower than the round-trip TC. While some MD biases are negative, the majority of the estimates grossly overestimate the values consumers place on the agritourism site. Possible reasons for any negative biases in the MD partial consumer surplus estimates in tables 3.7 and 3.8 could be based on: 1) different travel behaviors across regions; 2) further heterogeneity still not captured in the model; and/or, 3) inconsistencies in model seven parameter estimates from TC measurement error.

Analyzing the varying degrees of MD bias across regions and activities is one method to empirically highlight the importance of including heterogeneity in a travel cost model if more than one site is being valued. Excluding this regional or site attribute heterogeneity results in excluding important information of how consumer behavior differs across space and experiences. As one example, trips including the entertainment and events category of activities, which includes wine tours, is overestimated by 3,418% in Northern California while education is more modestly overestimated by 302% in the same region. This may indicate that if a recreational site or region is already implementing business strategies to extract a significant share of consumers' WTP through participation and recreation fees, such as the well-established wine-based activities in Sonoma and Napa Valley, then multi-destination travelers will most likely only deviate from their path of travel if the incremental travel cost is significantly low. Therefore, if this relatively small incremental distance for MD travelers is not measured, then large disparities are likely to exist between a respondent's incremental and cumulative travel costs. In other words, relatively expensive agritourism activities are likely to exhibit larger biases if MD status is not accounted for than less expensive activities, ceteris paribus.

Alternatively, the Central and Northeast Plains region also experiences significant bias in MD consumer surplus estimates with an average bias of 484% across activities. Recall that the motivation to aggregate these two large regions was due to a lack of observations, which can be viewed as a reflection of the low population density (and travel draw) of the aggregated region. The large distances between points of attraction, urban or natural, suggests multi-destination travelers participating in agritourism in this region are likely to travel farther distances from their point of origin and their destinations. This larger cumulative travel cost, therefore, inherently contains a larger sunk cost of travel when visiting a secondary or incidental agritourism site, leading to a larger disparity between the incremental and cumulative travel cost. Consequently, travel cost studies on recreation sites in more rural areas with larger travel distances are likely to suffer significantly more from not addressing the MD bias than those in more populated areas, ceteris paribus.

The aggregate CS estimate for agritourism in the Western US was calculated by weighting all partial CS values by the share of the activity in the region. The weighted CS's were then multiplied by the share of MD and PP travelers in the region. Finally, the weighted CS estimates were multiplied by 8% (the survey incidence rate) of each county's 18 years and older population for each region and aggregated across regions. When correcting for the MD bias through the use of incremental travel costs, the total CS of agritourism in the Western US is estimated to be $613.9 million, however the total CS falls to $466.9 million when the correction is not taken. This implies a -24% bias (-$250.6 million) from failing to ignore the sunk costs of travel. This unexpected negative bias is driven by, 1) the relative magnitudes of the PP and MD consumer surplus, 2) the shares of MD visitors in each region and activity, and 3) the popularity of each activity in a given region. If a region has a small share of MD visitors, even a large positive MD bias will be muted and possibly overcome by a relatively smaller negative bias in the PP consumer surplus estimates, which are generally much larger than the MD consumer surplus estimates. Similarly, some activities have very large positive MD biases, but if that activity is a small share of all activities in the region, then this bias will have little effect on the region's total CS estimate, and again, may be masked or overcome by smaller negative biases in the PP consumer surplus estimates.

In short, the varying degrees of biases in CS estimates and the overall downward bias in total CS implies a researcher cannot safely assume a specific site to have a positive or negative MD bias without considering a broader set of factors (such as the share of MD visitors, the relative explicit cost of the recreation experience, and the rurality of the surrounding area or distance to complement/substitute sites). Having this information may allow the researcher to assume a direction of the bias, but unfortunately, the magnitude of the bias likely varies across industries and sites, and therefore, cannot be estimated unless the incremental travel costs are directly collected and/or measured with the data collection instrument.

The above analysis makes significant contributions to the methodological TCM literature, however the correctly specified model (model four) also offers significant policy relevant applications. Farmers and ranchers either operating or interested in adopting an agritourism enterprise may benefit from not only knowing the CS of activities in their region, but also the implied price sensitivity of those activities for that specific region. CS estimates are useful for understanding the average willingness to pay by consumers, but changing the price of an agritourism activity to capture more of the CS also affects the number of visitors willing to pay for the activity. To capture this price sensitivity, the partial own-price elasticity of demand was estimated for each of the activities and regions, which in this context, is defined as the percent change in the number of visitors to an agritourism site from a one percent change in price for a certain region or activity.

Estimates from the truncated negative binomial travel cost demand model can easily be transformed into own price elasticities by multiplying the coefficient by the mean of the relevant variable (Grogger & Carson, 1991). These results are presented in table 3.9. The benefit of estimating part-worth price elasticities is based on the assumption that elasticities for activities are likely to vary across regions depending on consumer preferences, types of agritourism, and substitute recreational experiences in the region (Phaneuf & Smith, 2005). Given this information, policy makers, extension agents, and local tourism practitioners can determine which activities in their region are relatively more price sensitive than others, and to better inform agritourism operators on pricing strategies.

Table 3.9 Partial Own Price Elasticities

Activity/Region	**PP**	**MD**
Direct to Consumer	0.0219	0.0245
Events & Entertainment	0.1126	0.0440
Outdoor Recreation	-0.0336	0.0277
Education	0.0243	-0.0004
Mountain	-0.0226	-0.0364
Southwest	-0.0335	-0.0421
Central/Northern Plains	0.0127	-0.0074
Northern California	-0.0457	0.0048
Northwest	-0.0061	-0.0113
Texas	-0.0191	-0.0303

All elasticities are calculated at the variable's means and with model eight's results

The partial own-price elasticities in table 3.9 suggest that agritourism trips are relatively inelastic recreational experiences, however differences still exist within regions and activities. Some regions exhibit relatively flatter demand curves such as the Central/Northern Plains region for PP travelers and Northern California for MD travelers. The relatively flatter demand curves in these regions could be based on a number of factors. For one, these regions are either very remote, implying a large travel cost for any type of trip, or for others, be relatively more expensive in terms of paid activities, indicating that PP travelers in the Central/Northern Plains region and MD travelers in the Northern California region may see agritourism in these regions as luxury experiences relative to those in other regions. The elasticities for activities show a similar story. More expensive activities with perhaps a greater set of alternative recreational substitutes in the broader leisure sector, such as entertainment and events are relatively more elastic compared to potentially less expensive activities with less perfect substitutes like outdoor recreation in less congested, pristine agricultural areas. As another example, we can recall that hunting is a growing outdoor activity in the Mountain and Southern Plains, but is a relatively expensive recreational activity, suggesting that it might be more price elastic, however, there are few recreational substitutes that offer the same experience.

Conclusions

The travel cost method is one of the primary tools among non-market valuation techniques and has been used to influence policy decisions across multiple industries, traveler types, and regions around the world for several decades. The MD bias has posed a significant challenge in the travel cost literature since the 1980's because it was hypothesized that the bias could grossly overestimate the value of a site or activity and lead to suboptimal policy decisions. It is likely that the nature of tourism economics will only become more complex as modes of transportation, resources to plan travel independently, and global incomes rise, all potentially leading travelers to choose to add multiple experiences and sites to their travel plans. This paper shows how a unique flexible form that accounts for traveler, region, and activity heterogeneity, can be utilized to measure the MD bias and identify some policy implications for failing to account for the incremental costs of travel.

Similar to previous travel cost studies, this research finds that failing to account for the MD bias can lead to overestimating consumer surplus estimates for MD travelers. However, this study builds on previous research by revealing how MD bias can cause the demand parameters for PP travelers to be bias, confounding the total effect of the bias on total CS estimates. These results show that, on average, consumer surplus estimates for MD travelers are inflated by 244%, while PP consumer surplus estimates are underestimated by 10% if estimated without any effort to account for the bias. Despite the direction and magnitudes of these biases, when these consumer surplus estimates are scaled up to calculate the total consumer benefit of agritourism in the West the total bias is -24%. This result contrasts with previous travel cost studies that mostly assumed, or empirically found through other methods, positive biases in CS estimates. After exploring relationships between partial consumer surplus estimates, the counterintuitive result of the negative overall bias is likely a result of the travel cost measurement error producing bias estimates of PP demand parameters, highlighting the importance of accounting for consumer, site, and regional heterogeneity in future travel cost studies when using the TCM to value more than one site, which is commonly done in zonal TCMs.

A key finding is that the heterogeneity in consumer behavior significantly impacts the direction and magnitude of the MD bias. More expensive agritourism activities, such as wine tours in Sonoma county and Napa Valley, are likely to have much higher MD biases due to MD travelers being less willing to deviate longer distances from their primary path of travel to partake in expensive fee-based activities. When incremental distances for these travelers are not measured, significant disparities between incremental and comprehensive travel costs are likely to exist, and an activity's CS is likely to be grossly overestimated. Similarly, recreational activities in regions with lower population densities are likely to have largely inflated MD consumer surpluses due to higher cumulative travel costs relative to the incremental travel costs. In short, the MD bias is likely to be largest when the cost to participate in the activity is relatively high or if the activity is in a more remote location – implying larger comprehensive travel costs.

Due to the cost of primary data, the majority of travel cost studies, especially those using the zonal approach, are currently estimated using secondary data. Therefore, more research should be conducted using the incremental travel cost approach to develop upper and lower bounds for CS estimates produced by studies unable to comprehensively account for the bias. One suggestion for researcher who may not have information on incremental travel costs would be to estimate PP and MD visitors separately to protect the PP estimates from spillover biases from the measurement error in the MD visitors. This segregation may require a larger sample, but at least the researcher can be more confident in assuming an upward bias in the total CS of the recreation experience and can use heuristics based on the relative cost and rurality of the experience to understand the limits of their CS estimates.

Additionally, this study only calculates the bias using one model – a derivate of the Parsons and Wilson model. Other modeling approaches, such as the system of demand equations for bundled destinations by Mendelsohn, et al. (1992), should be compared to the incremental travel cost method to identify if one method of correction provides better estimates over others, specifically, a method that does not require primary data. Finally, there are likely other aspects of consumer behavior that affect the direction and magnitude of the MD bias, and exploring these facets may assist researchers in interpreting future and past travel cost models that do not comprehensively account for potential biases. Considering the large role the TCM plays in recreational, rural development, and resource management policy decisions, expanding the incremental TC literature in these directions will increase the validity of the TCM as a tool for calculating consumer welfare and identifying socially efficient outcomes.

CHAPTER FOUR: THE IMPACT OF CONSUMER AND LOCATION HETEROGENEITY IN DETERMINING TRAVELER BEHAVIOR

Introduction

In previous chapters I discussed and dissected some of the heterogeneity in the agritourism industry across regions, agricultural production types, and consumers. For the large part, I have found that there is potential for agritourism operators to exploit this heterogeneity by integrating the unique place-based and community factors nearby into their operational and marketing strategies, partly by recognizing and catering their activities to the type of travelers passing through their region. Hence, there is already evidence that agritourism is a flexible diversification strategy that can be implemented into a diverse set of agricultural production types in most regions of the U.S., albeit with different approaches and business models. However, a significant piece of heterogeneity that has not yet been explored, by myself or the literature, is the influence of site and consumer locational attributes on consumers' WTP for agritourism experiences. This further exploration has clear complementarities to the factors already considered, but allow for a more nuanced consideration of market potential.

In the interest of identifying a well-defined set of best practices for agritourism industry stakeholders to follow, producers would benefit from understanding the heterogeneity of consumer preferences, not only across market segments, but also across geographies. This piece of information is particularly useful for experience goods like agritourism, and other types of recreation that are location specific, because consumer preferences likely vary depending on the surrounding place-based attributes of the site. This is in contrast to most consumer goods that may vary in a finite set of product qualities, but whose general product type fundamentally remains consistent in the market, independent of site of purchase or consumption. Furthermore, it is not difficult to believe that a consumer's demand for an agritourism experience is not only dependent on the agritourism site's location, but also on the consumer's choice of where they live and its surrounding environment. Specifically, whether the consumer is a native or migrant to their location of residence, the choice to live in a specific area likely reveals something about their preferences for place-based and community factors. Assuming these revealed preferences exist, and are a product of the consumer's location of residence, we may deduce that a consumer's home locational attributes influence their demand for recreational experiences, such as agritourism. And, since marketing and promotion can be geographically targeted, this has management and marketing implications for the sector.

Given this discussion, this chapter applies data from a choice experiment to address the questions: how is heterogeneity in consumer preferences influenced by the consumers' location of origin, and on the supply side, how does WTP for specific agritourism qualities differ depending on the location of the agritourism activity? Exploring these joint research questions will enable agritourism industry stakeholders to not only identify relevant potential market segments, but also more efficiently target marketing efforts in areas of the contiguous U.S. where there is relatively higher demand.

Studying the influence that both site and consumer location may have on consumer preferences is particularly germane in the context of agritourism where the recreational activities many times take place in rural areas. Here, the access to urban or natural amenities is relatively more important than for urban recreational options due to the relatively higher travel costs incurred on any trip to more outlying locations. Past research has observed the importance of place-based factors in determining the success of agritourism (e.g. Van Sandt et al, 2018 and previous chapters), and some discrete choice experiments on recreational activities have explored consumer preferences for distances traveled to the recreational site, but no research has brought these two pieces together and investigated the roles consumer and site location play in choice behavior.

Previous studies have explored agritourists' WTP using revealed preference models (Carpio et al, 2008; Hill et al, 2014), motivations to adopt (Nickerson et al., 2001; McGehee & Kim, 2004; Barbieri & Mahoney, 2009; Tew & Barbieri, 2012; Barbieri, 2013), and spatial hot spots of agritourism incidence (Van Sandt et al, 2018). This study expands on the literature by exploring potential geographies and market segments that may be tapped to promote more efficient growth in the agritourism sector by capturing subtle differences in consumer valuation as it relates to site heterogeneity through a carefully structured choice experiment. Discrete choice experiments have been applied for decades to elicit consumers' willingness to pay and study their choice making behavior using a random utility framework originally developed by McFadden (1974).

Within the tourism literature, choice experiments have been used frequently to examine and forecast consumers' behavior given destination amenities, prices, and other qualities specific to the associated research topic. For example, Morley (1994) examines how price changes in transportation, hotel tariffs, and exchange rates have on tourists from Kuala Lumpur to eight different international cities. Morley (1994) finds that changes in transportation prices have the largest effect on consumers' choice to travel, perhaps implying that travel distances and the potential for multi-destination trips may be significant drivers of consumers' choice to visit an agritourism site. In addition, consumers also value destinations with scenic surroundings. In a consumer preference study on rock climbing in Scotland, Hanley, Wright, & Koop (2002) find that the scenic quality of the surrounding area of a climbing route significantly adds to the respondent's value of the activity, but only include four levels of the attribute ranging from "not at all scenic," to "very scenic." Considering the value consumers placed on scenic landscapes, understanding which "scenic" qualities consumers value for a particular recreation site may be of great interest for recreation managers and environmental policy assessments

If I may slightly digress here, travel distances and the opportunity to visit multiple sites on a single trip may play an important role in assessing the agritourism sector's potential as a rural development tool. Several sources (such as Wilson, Thilmany, & Sullins (2006), Bagi and Reeder (2012a), and others) note agritourism's ability to stimulate economic growth, however, Brown et al. (2014) found that community-focused agriculture (defined as direct to consumer sales and agritourism) did not have a significant impact on economic growth at a national level, but did have some varying effects across regions. These mixed findings across regions and studies are likely an indication of the confounding role of heterogeneity in both consumer preferences across regional markets and in the locations of local agritourism sites relative to other urban and natural amenities. In the context of this chapter, if consumers significantly value being in close proximity to natural or urban attractions, this may be an indication of their preference for multi-destination trips and willingness to spend dollars at other local attractions when on an agritourism trip. This could therefore imply that some agritourism sites near other recreation opportunities may be more conducive to increasing farm incomes and stimulating rural development, due to their greater draw for agritourists looking to minimize travel costs. Travelers participating in multiple recreational activities in a localized region will likely spend more time in the area and in surrounding local communities, increasing the potential for them to spend dollars at other travel related businesses (i.e. restaurants, hotels, shops, outfitters, etc.).

Other more remote sites may instead have to cater to the preferences of their surrounding populations, and may therefore not benefit from the inflow of outside dollars that are seen as one of the potential drivers of rural development.

It is also possible that a consumer's experience in a specific recreational activity may affect their WTP for certain attributes of the trip, including their WTP for familiar or unique activities and locational attributes. Hynes, Hanley, & Scarpa (2008) find experience in kayaking positively impacts consumers' WTP for additional levels of star ratings, water quality, and the scenic surroundings. While this provides evidence of consumers' preferences being contingent on their experience with the recreational activity, it stops short of considering other potential sources of consumer preference heterogeneity such as consumer sociodemographics, location of origin, and exposure to natural/urban amenities. Given all outdoor recreation experiences are influenced by location and place-based factors, understanding the preferences of the local population surrounding a particular outdoor recreation site can provide producers and tourism development practitioners with more accurate part worth values in a highly heterogeneous industry. Furthermore, knowing the location of distinct market segments for a specific agritourism activity and augmenting this with existing state tourism data, may provide industry stakeholders with more precise estimates of market size and more efficient marketing strategies.

Juutinen et al. (2011) use a choice experiment to explore management policies in Finnish national parks and find consumers experience welfare gains from greater biodiversity and less manmade structures on the trails. These studies provide some context of how to approach the U.S. agritourism sector, but a more targeted look at issues of key interest to the sector are warranted. For example, if Finnish national park goers place value on scenic qualities for ecotourism then one may also be able to infer that, for U.S. agritourism, more remote areas with more rustic lodging and outdoor activities could elicit a higher WTP for at least some subsets of consumers.

As the reader can probably deduce from the definition of agritourism presented in the previous chapters, for the USDA and this study's scope, agritourism can take many different forms such as wineries, corn mazes, food preparation/preservation classes, hunting, photography classes, farm dinners, and farm and ranch heritage experiences. It is this generality and the ability to cater an agritourism enterprise to the farm or ranch's historical or amenity assets and other strengths that makes agritourism a challenging industry to provide part-worth site quality values for. Considering this site and consumer heterogeneity and the multiple objectives of this study, I integrate word of mouth recommendations (online review metrics), place-based factors, and activity choices in a choice experiment to jointly assess the value a set of travelers place on various aspects of an overnight agritourism experience.

More specifically, utilizing a latent class logit model I estimate the WTP values for the agritourism attributes and include county level class membership covariates to capture the role an individual's environment has on their choice behavior. Since the class membership variables are measured at the county level, I am then able to predict the membership shares for most U.S. counties and use a geostatistical tool called empirical Bayesian kriging (EBK) to create a continuous map of relative preferences for six different agritourism attributes across the contiguous U.S. Results indicate that consumers' exposure to natural amenities, experience with agritourism, and sociodemographics significantly impact the values they place on an agritourism site's activities, lodging, surrounding area, and online reviews. Furthermore, the relative WTP maps show that the heterogeneity in consumer preferences across the contiguous U.S. does not follow any specific dichotomy (e.g. rural-urban) implying that market shares are contingent on a complex portfolio of place-based and community factors.

The rest of the chapter is organized by first reviewing the experimental design and the model estimation processes. After presenting part-worth values of site attributes and class membership coefficients, I predict differences in WTP values across locales not observed in the sample using Empirical Bayesian Kriging. Given some limitations in available survey and county level data, this method provides a complete visual representation of the spatial distribution of WTP values across the contiguous U.S., thereby allowing me to more thoroughly examine market potential and opportunities based on the unique place-based and community characteristics of a particular location. Finally, I conclude with a discussion and summary of the results, areas for future research, and policy implications.

Methods

Survey Data

As mentioned in chapter two, one survey instrument was used to collect data for multiple research questions and studies including the travel cost model and the choice experiment presented in this chapter. To review, the survey data was collected in April of 2015 through Taylor Nelson Sofres' (TNS), an online survey company, using their pre-recruited panels. The sample contains a total of 1,501 respondents from all over the United States. Of these 1,501 respondents, 1,001 visited an agritourism site in the western U.S. during the previous twelve months (with 8% incidence in the broader sample), while the other 500 did not participate in an agritourism activity but had traveled out of the state in the past twelve months (with 53% incidence in the broader sample). The first subset of respondents, the agritourists, were collected specifically for the travel cost method as it examines revealed preference, while the second subset of respondents was collected in order to expand our sample to a broader population for the choice experiment and explore the market potential one might choose to explore through stated preferences. While only the 1,001 agritourist respondents were used in the TCM, all 1,501 respondents were used in the choice experiment.

The choice experiment occurred at the end of the survey and consisted of two different experiments (a “day trip” and a “ranch-stay”) that were randomly and evenly divided in each subsample (regional and national). Using a stated choice software, Ngene, I adopted a sequential orthogonal design for the choice experiment. Each choice experiment was divided into four randomly assigned blocks with nine choice situations in each block and two alternatives per choice situation. The entire survey was estimated to take approximately 15 minutes to complete. As the reader can deduce, the sample is somewhat biased toward the Western U.S. given 66% of the sample visited an agritourism site in the Western U.S., so results should be interpreted keeping this in mind. However, it should be repeated that respondents in the agritourism subsample were not selected based on where they came from but rather where they participated in an agritourism experience, meaning that the preferences displayed in this choice experiment analysis are not necessarily geographically biased in terms of the respondent’s origin. A map of respondents’ origin of residences can be found in the appendix (figure A.9).

The choice to administer the survey online through a research company is appropriate for two reasons. First, internet survey companies that offer incentives to a pre-recruited sample tend to have higher response rates, and second, a greater sense of authenticity may be developed if the individuals, who select a tourism destination through a simulated travel website called Trip Guru, actually make their decisions online in front of a computer. While I can find no studies on how the design and presentation of a tourism based choice experiment affects the respondent’s decisions, this is a well-studied topic in the broader choice experiment literature. The design of the experiment plays an essential role in developing a choice experiment, but as far as different choice experiment designs (i.e. number of levels, situations, etc.), Caussade, Glaeser, & Kerr (2005) find that while varying the design of their choice experiment led to greater choice inconsistencies it did not significantly impact WTP estimates. In an attempt to present a more authentic experiment Bateman, Day, Jones, & Jude (2009) go as far as to compare numeric and visual representations of information using virtual reality technology. Bateman et al. find significantly lower asymmetry between WTP and willingness to accept values when using the virtual reality technology, compared to just the numerical information, however, I did not have access or funds to replicate a similar method of analysis that would assess this sense of authenticity. Regardless, the main goal of the researcher should be to simulate reality, and fortunately, the internet has taken a large role in guiding consumers' travel decisions, making it a justifiable and valid mode in which to present information for a choice experiment on agritourism.

The analysis in this chapter only applies to the data collected for one of the trip choices: namely, the overnight ranch-stay agritourism experience. The attributes included in the choice situations differed depending on which alternative the individual was assigned to and each attribute took different levels across choice situations. Varying attributes for the ranch-stay, defined as a five-night package for two people, were price, estimated travel cost from home, lodging options, a "star-rating" from previous travelers' online reviews, activities included in the package, attributes of the surrounding area, and distance to a national park.[11] The attribute levels for the overnight ranch-stay are presented in table 4.1. The prices of the agritourism experience

[11] The day trip travel listing was for one individual and included the following attributes: price, distance from home, rating from online reviews, and activities included in the package.

Table 4.1 Choice Experiment Attributes and Levels for a "Ranch-Stay" Agritourism Trip				
Attribute	Level 1	Level 2	Level 3	Level 4
Price	$500	$1,500	$2,500	$3,500
Star Rating	3.5 stars	4.5 stars		
Lodging	Tent/yurt campsite with fire pit (access to vault toilet facilities)	Rustic ranch house (electricity, private bath, and kitchen)	Modern cabins (with electricity and private bath)	Bed and Breakfast with semi-private bath
Activities	Historical Excursions (crafts, food, and tours)	Horseback trail rides		
Surrounding Area	Less than an hour from the city, with diverse restaurants and entertainment options	Located minutes from a historical main street with unique local dining and drinking options	Located in a scenic and remote area	
Distance from National Park Entrance	30 minutes	One hour	One hour and 30 minutes	Two hours
Constant attributes across all alternatives: five-night package for two people, 153 reviews/ratings, estimated travel cost from home: half a day's travel with flight for $450, and fishing and photography hikes/classes. Sequential Orthogonal Choice Experiment Design Summary: 750 respondents, four blocks, nine choice situations per respondent, two alternatives per choice situation, 6,750 total observations				

were crosschecked by an advisory board containing industry stakeholders including agritourism operators, state tourism practitioners, and agritourism association professionals. The price attribute levels were less uniformly distributed at the endpoints and were the only attribute that was not completely uniformly distributed.

Theory and Empirical Model

A random utility framework was adopted to model the travelers' behavior in choosing one agritourism site over another. Under the assumption that individuals are utility maximizing agents, an individual n will select agritourism alternative i over j alternatives only if alternative i offers them the most utility relative to alternative j.

(4.1) $$U_{ni} > U_{nj}$$

Given any choice situation, the researcher only observes information on the alternatives' attributes z_{nj} and some information about the individual themselves r_n, resulting in an observed portion of the individual's utility function $V_{ni}(z_{ni}, r_n)$, and an unobserved portion ε_{ni}. The probability an individual chooses alternative i over j alternatives, denoted P_{ni}, then becomes:

$$P_{ni} = Prob(U_{ni} > U_{nj})$$

(4.2) $$P_{ni} = Prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj})$$

$$P_{ni} = Prob(V_{ni} - V_{nj} > \varepsilon_{nj} - \varepsilon_{ni})$$

Where the probability that an individual chooses i is the probability that the difference between the observed portions of utility is greater than the difference between the unobserved components. The cumulative probability is the integral over the density of the unobserved portion of utility, $f(\varepsilon_n)$:

(4.3) $$P_{ni} = \int_{\varepsilon} I\big(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \;\; \forall j \neq i\big) f(\varepsilon_n)\, d\varepsilon_n$$

If we assume the errors are independent from irrelevant alternatives (iia) and that the unobserved component of the respondent's utility function is Gumbel distributed then Train (2008) shows that a logit choice probability can be obtained from this cumulative distribution:

(4.4) $$P_{ni} = \frac{\exp(\beta' z_{ni})}{\sum_j \exp(\beta' z_{nj})}$$

where the probability of individual n choosing alternative i is a function of z independent variables observed by the researcher with preference parameters β to be estimated.

While limited dependent empirical models such as the multinomial logit are popular due to their ease of interpretation and closed form, the strong assumption of the individuals' errors being iia has led to some researchers choosing more flexible forms such as the mixed-logit that allows the researcher to specify continuous joint distributions for the unobserved errors (Train, 2008).

Unlike the models described above, the latent class logit model assumes that consumers can be classified into a discrete number of market segments, or classes ($s = 1, \dots, S$) based on their preference similarities and differences, allowing the model to endogenously describe consumer heterogeneity rather than it being assumed to follow some distribution specified by the researcher. Since I have no reason to believe that agritourists behave similarly across the U.S., I adopt the latent class logit model over the alternatives to avoid biased estimates that may arise from the assumption of homogeneity across consumers. One last reason I chose the latent class logit over model alternatives was the greater flexibility it offered given the main effects design of the choice experiment.

Given some number of classes S set by the researcher, the probability that individual n is observed in class s is:

$$\gamma_{ns}(\varphi_s) = \frac{\exp(\varphi_s x_n)}{1 + \sum_{s=1}^{S-1} \exp(\varphi_s x_n)} \tag{4.5}$$

where φ are coefficients of individual specific covariates x_n that determine class membership, and class S is dropped for identification (Pacifico & Yoo, 2013). The probability of individual n choosing alternative i over alternative j with T choice occasions, is the sum of each class membership probability multiplied by the conditional probabilities of individual n choosing alternative i conditional on being in class s across all alternatives and choice occasions:

$$P_n(J = i|z_n) = \sum_{s=1}^{S} \gamma_{ns}(\varphi_s) \prod_{t=1}^{T} \prod_{i=1}^{J} \left(\frac{\exp(\beta_s' z_{ni})}{\sum_j \exp(\beta_s' z_{nj})} \right)^{C_{nit}} \tag{4.6}$$

where C is an indicator function that takes on a value of one if individual n selects alternative i on occasion t and zero if otherwise (Pacifico & Yoo, 2013). Note the conditional probability that individual n chooses alternative i given they are in class s, the right-hand term in parentheses, is simply the multinomial logit specification for a respondent in class s (Shen, 2009). To summarize, the objective is to estimate coefficients β and φ given each alternative's attributes z and individuals' characteristics x. This information will allow me to control for individual heterogeneity while determining the influence location has on individuals' choice behavior and WTP for each attribute across classes.

Given difficulties in empirically estimating the log likelihood function with the standard Newton Raphson search algorithm, the Expectation-Maximization (EM) method has been proposed and found to be superior in run time and allows for more flexible approximations of the parameter distributions (Bhat, 1997; Train, 2008). I estimate the latent class logit model in STATA following Pacifico & Yoo's (2013) command *lclogit* that makes use of the EM method, and then use STATA's *gllamm* to estimate the standard errors of the parameters.

Several previous studies perform various tests to determine the number of latent classes that captures the bulk of consumer heterogeneity without overfitting the model. While studies such as Nylund, Asparouhov, & Muthen (2007) find bootstrapping methods to be superior in determining class enumeration, the Bayesian Information Criterion (BIC) seems to be more commonly used (or at least traditionally used) due to its comparable results and relative simplicity. I chose to use the BIC option in determining class enumeration, selecting the model with the smallest BIC.

Latent class models have been used extensively in the literature to calculate WTP values across market segments (see Hidrue, Parsons, Kempton, & Gardner, 2011; Hensher & Greene, 2010; Wen & Lai, 2010; and others), but few studies have taken advantage of the latent class model's preference heterogeneity across individuals to examine preference heterogeneity across space. One study that follows methods similar to my own is Brouwer, Martin-Ortega, & Berbel (2010) where the authors model preference heterogeneity for water quality throughout a river basin in southern Spain. The authors include each respondent's residence of origin as a covariate in determining class membership to examine preference heterogeneity across space and find that economic welfare measures change across space depending on environmental conditions and socioeconomic demographics where they reside (Brouwer et al, 2010).

Similar to Brouwer et al. (2010), I include a set of five covariates to determine class membership, all of which are measures based on the respondent's county of origin. The hypothesis here is that an individual's choices are partly driven by their daily environment. For example, a respondent from the Rocky Mountains may have different recreation preferences than someone from the Northern Plains states due to their revealed preference to live in a region where natural amenities are readily available. The explanatory variables that are included to determine class membership, x_n in equation 4.5, are agritourism revenue, per capita income, minutes to a population center of at least 50,000 people, the USDA's natural amenities index, and the respondent's age which was compared to the county's median age during the spatial interpolation process. Since these measures are available for most counties, I am able to estimate each county's WTP by first estimating each county's class membership probability and weighting the WTP estimates within classes to capture the heterogeneity of preferences within the county. Essentially, I am using the respondents to determine market segments and then apply the counties' characteristics to estimate the WTP for an average consumer in each county where data is available. This method allows me to identify what locational factors influence consumers' choice behavior and explore geographical markets where the WTP for a specific activity or agritourism characteristic is significantly high relative to others.

While Campbell (2007) take a very similar approach to mine, estimating a mixed logit and using kriging to predict WTP for rural landscape improvements in the Republic of Ireland, this study differs in its the treatment of preference heterogeneity. While Campbell (2007) use a mixed logit, I capture consumer heterogeneity through a latent class logit model using origin specific information to determine the differences in WTP across space. Further, I also use the county level information to determine the weighted average WTP of an area rather than the individual respondents' WTP. Given the geographical scope of this study, the contiguous U.S., and the project's objectives, I believe my methods of basing an area's WTP on aggregate county data rather than the characteristics of a single respondent to be more appropriate for estimating market potential.

Spatial Interpolation

The objective of the spatial interpolation of relative WTP estimates across the contiguous U.S. is two-fold: 1) to observe the spatial distributions of preferences to provide insight on when and how locational factors might influence preferences; and 2) to provide agritourism industry stakeholders with information on geographical markets with relatively higher or lower WTP values for specific agritourism attributes to improve the efficacy of marketing efforts. One method would be to predict each respondent's WTP and mapped these to show the spatial distribution of preferences for each attribute across space. However, the number of unique counties of origin in the sample only account for less than one sixth of all counties in the contiguous U.S., making for an incomplete map of preferences (a map of respondents' counties of origin can be found in the appendix: figure A.9). If the average max posterior probability in the latent class analysis is significantly high, indicating the model preforms relatively well at predicting respondents' class membership, then a better method would be to utilize the county level nature of the membership covariates to predict class membership for each county and weight WTP values accordingly.

To illustrate this process further, I first used the estimated class membership coefficients to predict the class membership probabilities for each county with available data. Since the class membership probabilities always sum to one, I then used these predicted class probabilities to weight the WTP estimates for a specific attribute for each county, resulting in a weighted WTP value for the attribute of interest. However, some counties have undisclosed agritourism data meaning not all counties' WTP for the attributes can be predicted. Furthermore, counties vary significantly in geographical size with counties in the West being on average four times as large as those in the eastern U.S. (not including Hawaii or Alaska). These two issues lead to maps that are both incomplete and misleading when identifying the location of potential market opportunities.

To circumvent this issue, I adopt a geostatistical interpolation method called kriging that effectively exploits the spatial autocorrelation between estimated points to interpolate values in between each county's population centroids. There are several methods for interpolating un-sampled points that can more or less be broken down into deterministic and probabilistic methods. Deterministic methods, such as inverse distance weighting, use a predetermined function of distance between known points to interpolate the values of unknown points. Probabilistic methods, such as kriging, apply statistical measures to not only interpolate un-sampled points but also provide a measure of uncertainty around such points (Krivoruchko, 2012). In addition to providing a method for ascertaining the WTP for specific attributes in counties with undisclosed data, it also generates a more continuous depiction of WTP values across space reducing the concerns with the issues noted above. A greater explanation of the EBK process and summary statistics from the analysis can be found in the appendix.

Empirical Results

Using the BIC I determined that four latent classes captured the majority of consumer heterogeneity. The model converged after 67 iterations using the EM method and had an average max posterior probability of 88%, indicating that the model was reasonably successful in modeling the heterogeneity of consumers' preferences. As a reminder, when discussing the significance and sign of the class membership coefficients in the following paragraphs, these interpretations are relative differences when compared to the reference group, Class four. The latent class model estimates can be found in table 4.2.

Given the WTP estimates and class membership coefficients that were statistically significant at the 10% level (or greater), I labeled the four classes: Comfort Tourists (24.7%), Adventurers (33.8%), Road Trippers (20.7%), and Urbanites (20.8%). Class four, the Urbanites, were the reference group for identifying class membership coefficients. Comfort tourists likely came from counties with relatively high incomes per capita, are relatively price sensitive, and prefer more upscale lodging options like the ranch house option ($185.77). The Adventurers had a strong preference for tent lodging ($1,316.33) and horseback riding ($1,136.17), and likely came from counties with relatively high agritourism activity, income per capita, and natural amenities, while being relatively younger individuals. The Road Trippers tended to be older and from areas with fewer natural amenities. The Road Trippers preferred more upscale lodging options (cabin: $3,695.72), historical excursions ($513.93), a close proximity a small town ($1,826.90), and close proximities to National Parks (a discount of $496.20 per 30 min. of drive time). Finally, the Urbanites were relatively price sensitive and had a strong preference for urban amenities ($2,624.39), quality lodging (cabin: $185.51), and historical excursions ($1,449.97) instead of horseback riding.

Surprisingly, the travel time (in minutes) to an urban area, defined as a city with a population of at least 50,000, did not significantly influence how a consumer was categorized into a market segment for this study. While this variable may still have an effect on some consumers' preferences, it does not appear that it had explanatory power in parsing out consumer preference heterogeneity in this sample. Both the Road Trippers and the Urbanites had comparably strong preferences for ranch-stay agritourism sites that were also nearby urban or small-town amenities, but these characteristics were insignificant for the Comfort Tourists and Adventurers.

As hypothesized, the natural amenities index had an inverse relationship with classes preferring more upscale lodging options and outdoor activities, thereby indicating that a respondent's surrounding environment can impact their recreation preferences. The Adventurers were from relatively more natural amenity rich areas and valued camping and horseback riding. In contrast, the Road Trippers who were from areas with fewer natural amenities, had comparatively stronger preferences for more upscale lodging and historical excursions. Surprisingly the Adventurers did not have a significant preference for being close to National Park entrances but the Road Trippers and Urbanites did value this attribute in the context of an agritourism experience.

Living in counties with higher incomes per capita had less of an effect on consumers' price sensitivity, and instead, the relative amount of agritourism activity in the respondent's home county drives differences among traveler preferences. Both the Comfort Tourists and Adventurers came from relatively wealthy areas, but the Comfort Tourists were much more price sensitive than the Adventurers who had relatively high WTP values. However, the Adventurers and Road Trippers, who are both from counties with more agritourism activity, were relatively less price sensitive than the Comfort Tourists and Urbanites, both of which had relatively less agritourism activity around their residence of origin. In other words, consumers with greater familiarity with agritourism had greater WTP values for agritourism attributes.

The Adventurers, who are younger and assumedly more tech savvy, put the highest value on an additional star rating from the simulated travel site. Yet, Road Trippers, who are decidedly older than the Adventurers, also placed a relatively high value on better online ratings. Due to the fact that both of these classes are surrounded by relatively high levels of agritourism activity, these respondents may be more familiar with agritourism and seeking well-reviewed agritourism experiences to complement what is available to them in their home region. While consumers who are more familiar with agritourism appear to be more discriminating of the value of an online star rating, producers should be aware the Comfort Tourists and Urbanites also value higher ratings, underlining the somewhat obvious importance of providing an enjoyable experience for visitors, and perhaps incentivizing them to offer up reviews, so that their reviews and online reputation can be a strategy to differentiate their operation.

Spatial Interpolation

Given the significance and magnitude of some of the class membership covariates, it seems that one's location of origin has a significant impact on consumers' choice behavior.

Table 4.2 Latent Class Logit

	Class shares	Class 1		Class 2		Class 3		Class 4	
		24.7%		33.8%		20.7%		20.8%	
	Variables	Comfort Tourists		Adventurers		Road Trippers		Urbanites	
		$\hat{\beta}$	$\widehat{WTP}$	$\hat{\beta}$	$\widehat{WTP}$	$\hat{\beta}$	$\widehat{WTP}$	$\hat{\beta}$	$\widehat{WTP}$
Lodging	Price	-0.0028***	-	-0.0002***	-	-0.0007***	-	-0.0029***	-
	Rating	0.2527*	$ 89.34	0.2290***	$ 1,254.58	0.3060***	$ 454.08	0.6704***	$ 231.97
	Tent	-0.1103	$ -39.00	0.2402*	$ 1,316.33	-2.1001***	$ -3,116.31	-7.2454***	$ -2,507.16
	Ranch House	0.5254*	$ 185.77	0.0999	$ 547.51	1.9278***	$ 2,860.73	-2.4342***	$ -842.31
	Cabin	0.1593	$ 56.31	0.1019	$ 558.42	2.4905***	$ 3,695.72	0.5361	$ 185.51
	Bed and Breakfast	-	-	-	-	-	-	-	-
Act.	Historical Excursion	0.6282***	$ 222.11	-0.1012	$ -554.53	0.3463**	$ 513.93	4.1903***	$ 1,449.97
	Horseback	0.4287	$ 151.58	0.2074*	$ 1,136.17	0.3134	$ 465.07	-7.9613***	$ -2,754.86
Surr. Area	City	-0.3617	$ -127.91	0.0220	$ 120.78	0.5845***	$ 867.40	7.5842***	$ 2.62
	Quaint Town	0.1401	$ 49.53	-0.1069	$ -585.91	1.2311***	$ 1,826.90	0.1170	$ 0.04
	Remote	-	-	-	-	-	-	-	-
	Distance to Nat. Park	-0.0462	$ -16.33	0.0126	$ 68.96	-0.3344***	$ -496.20	-1.5887***	$ -0.55
Class Membership	Agritourism Revenue	0.2146		0.2970**		0.2681*		-	
	Ln(income/capita)	0.6411***		1.0627***		0.2245		-	
	Age	0.0036		-0.0286***		0.0361***		-	
	Natural Amenities Scale	-0.0397		0.0740*		-0.1365***		-	
	Minutes to Urban Area	0.0041		0.0026		-0.0015		-	
	Constant	-2.5733***		-3.0390***		-2.3065***		-	

*** P-value<1%
** P-value<5%
* P-value<10%

Observations: 751
Average Max Posterior Probability: 0.8809

Furthermore, there seem to be some relatively strong preferences for agritourism attributes that are dependent on location such as camping, horseback riding, history, and of course the surrounding area and proximity to a National Park entrance. However, there still does not seem to be a clear narrative on how agritourism preferences vary across space, or where potential and untapped agritourism markets may exist. To expand on this empirical question, I use Empirical Bayesian Kriging (EBK) to interpolate the WTP for specific attributes across space. EBK allows the researcher to relax assumptions about some of the semivariogram features, such as the nugget, sill, and lag, by using an iterative process to find the best fit of the semivariogram that can then be used to better explain the spatial relationships between known points. Rather than mapping the predicted WTP values, I map the percent difference from the class weighted mean WTP for each county population centroid for six agritourism attributes. The mapped attributes were chosen based on their significance in the LCA as well as whether or not the variable was assumed to have a significant spatial component. For example, the values consumers place on star-ratings are presumably less likely to vary across space than the values they place on location dependent attributes, such as camping.

Figures 4.1 through 4.6 show maps of spatially distributed differences in WTP estimates for specific attributes weighted across a county's predicted class membership probabilities (i.e. shares). Predicted class membership shares were calculated using the class membership variables and county level data to calculate the log odds and eventually the class shares of consumers in each county. WTP values were then weighted by these class shares in an attempt to model the average traveler's preference in that county. Once again, to enhance the spatial interpolation method's validity, I used each county's population centroid as the point of information.

At first glance, it appears that some of the activities and lodging maps are practically mirror images of each other. This should not be surprising since they essentially represent opposite experiences for those attributes, or in the case of activities, substitute experiences. For example, horseback riding, a relatively physically demanding and outdoor experience, is much more popular among those who live in the West, Florida, and on the Northeast coast. The relatively higher WTP in these first two regions immediately makes sense since both of these areas have a long heritage surrounding cowboy and horse culture. In addition, relative to the Midwest, these three regions have relatively more natural amenities (see figure A.7 in appendix), presumably enhancing participants' utility derived from a wide range of outdoor activities.

Not all attributes show the same degree of spatial heterogeneity in consumer preferences for agritourism qualities. Some attributes such as being in close proximity to a city varied very little across those from different locales, while other attributes such as tent lodging and the activities varied significantly among these groups (see table 4.3). In general, WTP for attributes describing the surrounding area varied the least across space while outdoor activities and lodging varied the most. When agritourism qualities do not vary significantly across space, targeting specific geographical markets leads to smaller efficiency gains, however, this does not mean that there is no heterogeneity in preferences across individuals. For example, preferences for being in close proximity to a city only varied within 11.7% of the mean, however, class three, the Road Trippers, placed significant value on this particular attribute, indicating that there may still be benefits from targeting a specific market segment. In summary, not all qualities exhibit patterns of significant spatial heterogeneity in preferences, but consumers' WTP for some recreational qualities can still vary considerably across market segments when differentiated by characteristics other than home locale.

Table 4.3 Spatial Heterogeneity of WTP for Agritourism Attributes

Attributes	% Difference from Mean WTP	Relative WTP Interval Length	High WTP Regions	Low WTP Regions
Surrounding Area: City	8.86% to -2.84%	11.7	West, South, and NE Coast	Midwest and California Coast
Park Lands Proximity*	16.67% to -14.91%	31.58	Midwest	West
Ranch House Lodging	5.28% to -38.25%	43.53	Midwest, East Coast, and South	California
Historical Excursion	26.00% to -24.74%	50.74	Midwest	Pacific Coast and Southwest
Tent Lodging	39.48% to -25.48%	64.96	West	Midwest
Horseback Riding	24.60% to -52.17%	76.77	West, Florida, and NE Coast	Midwest

* Estimates for proximity to NPS lands been transformed so that a higher WTP is associated with being in closer proximity to NPS lands.

In summary, the spatial distributions of WTP for agritourism attributes suggests that some preferences are driven by the consumer's familiarity with the attribute, like horseback riding, while other preferences are driven by the relative uniqueness of the attribute compared to where they reside, such as the proximity to NPS lands. Region, natural amenities, and to some extent rurality all appear to be correlated with patterns in WTP across space. Yet, in some cases spatial patterns of WTP seemed to be less prevalent, or even inconsistent in regards to the place-based factors considered, implying that there could be latent factors influencing a region's WTP for some agritourism attributes.

Discussion and Conclusion

Many agricultural producers and economic development practitioners in rural areas are exploring opportunities to diversify farm- and ranch-based businesses and keep rural economies viable. Clearly these motivations to diversify an agricultural business and rural economy must have a strong grounding in where a farm or ranch is located and what they produce. Agritourism is a particularly attractive diversification strategy because activities can vary widely and take advantage of farms or ranches' assets, location and other strengths such as unique crops and livestock enterprises of interest to visitors or proximity to population centers or adjacent attractions or cities to visit. In addition, agritourism has been identified as an underutilized strategy for rural development by the USDA.

Given the important differentiation potential an agritourism enterprise can leverage based on their location, this chapter explores how a consumers' area of residence affects their choice behavior, and uses this information to identify spatial patterns in consumer preferences and potential agritourism markets. Using a choice experiment with varying qualities for an overnight ranch-stay agritourism experience and a random utility framework, I estimated consumers' WTP for said qualities while capturing preference heterogeneity using a latent class logit model. I established four classes that captured the bulk of the respondent heterogeneity and labeled these classes as Comfort Tourists, Adventurers, Road Trippers, and Urbanites. As indicated by the class names, particularly large differences in preferences existed related to a traveler's home (whether they come from urban surroundings) and interest in outdoor activities.

More so than the per capita income of the respondent's origin, existing agritourism activity had a large impact on the respondent's WTP, implying that those familiar with agritourism are more likely willing to pay more for an overnight agritourism experience than someone who has had less exposure to the industry and its offerings. Producers should interpret this result carefully. A dearth of agritourism activity in a county may be correlated with lower WTP values for agritourism activities, but the classes that were more likely from these areas still had higher WTP values for some specific activities that producers can still take advantage of. These results are somewhat echoed in the first chapter where the I find evidence of agglomeration economies in the U.S. agritourism industry resulting from information sharing or decreased marginal travel costs for multi-destination agritourists.

Taking the information gleaned from the latent class analysis, I utilized the class membership covariates to determine the class shares for each county and used these to weight the WTP estimates for various attributes in the choice experiment. These points of information were then used in an Empirical Bayesian Kriging analysis where the spatial relationships between points were used to interpolate relative differences in WTP values across the contiguous U.S. This contribution to the literature not only expands our knowledge of how consumer preferences vary across classes and space for an experience product, but also provides a valuable tool for future agritourism sector development by enabling targeted marketing strategies in terms of market segments and geographical markets.

The spatial patterns of WTP differences most notably showed a stronger preference for outdoor activities and more rustic lodging options when considering visitors living in the Western U.S., Florida, and on the NE coast. While urbanization of the respondent's origin did not significantly parse out differences in choice behavior, the spatial patterns show regions that strongly prefer, or did not prefer, urban amenities and the distance to a National Park entrance, though the variation in WTP values among these two attributes was significantly less relative to other attributes. Considering that distance traveled can be interpreted as an implicit cost to consumers, it is reasonable to believe that for many agritourism establishments the preferences of consumers in nearby areas play a greater role in the success of an enterprise than those same preferences averaged across the country.

The spatial distributions of WTP for agritourism attributes suggests that some preferences are driven by the consumer's familiarity with the attribute, like horseback riding, while other preferences are driven by the relative uniqueness of the attribute compared to where they reside, such as the proximity to NPS lands. Region, natural amenities, and to some extent rurality all appear to be correlated with patterns in WTP across space. Yet, in some cases spatial patterns of WTP seemed to be less prevalent, or even inconsistent in regards to the place-based factors considered, implying that there could be latent factors influencing a region's WTP for some agritourism attributes.

Some consumer's preferences appear to not only be related to the physical surroundings of an area, but also subcultures. Consumers in the Western U.S. and Florida display relatively strong preferences for horseback riding, which is in line with these two regions' robust heritages in cowboy culture. In contrast, overnight agritourism enterprise operators with historical excursions may be more efficient at marketing in the Midwest, a region with a relatively greater share of rural communities that principally rely on the agricultural sector. Other regions, such as the Southeast, seem to not have as strong of preferences for either of the two activities.

Across the contiguous U.S., WTP values for populations adjacent to NPS lands are either at or below the mean WTP for an agritourism experience in close proximity to NPS lands. The simplest explanation for this is that consumers who live near NPS lands have some of their demand for these recreational lands met by the local supply of public lands, thereby lowering their WTP for being near additional NPS lands during travel. Consumers who live farther from NPS lands tend to be willing to pay more for an agritourism experience in close proximity to a NPS land due to the lack of supply of such recreational lands near their location of residence. Comparing figure three to a map of driving times to NPS lands (figure A.1. in the appendix) reveals a picture consistent with this narrative. In short, consumers who live in natural amenity rich areas, such as the Pacific Coast, Rocky Mountains, and Florida, tend to value the proximity to NPS lands less compared to those who live in natural amenity poor areas (compare figure three to figure A.2. in the appendix).

In addition, this differences in preferences across market segments is evidence that while the Adventurers prefer outdoor lodging and activities, they are more often primary purpose travelers than the Road Trippers and Urbanites who would rather make agritourism one stop on a multi-destination trip that includes venues such as National Parks. Not surprisingly, these findings provide evidence that producers in natural amenity rich areas may have more success with agritourism models that integrate outdoor activities, especially if they are near National Park entrances, where they have the opportunity to attract the travelers who are similar to the Road Trippers and Urbanites. Further opportunities to market toward the preferences of the Road Trippers and Urbanites exist for agritourism operations near urban areas. These results highlight the potential for agritourism operations to generate more economic activity by offering activities that are consistent with the preferences of the market segment with the aligned locational preferences.

One caveat of this work is the fact that the survey instrument was designed for several other objectives within an umbrella projects which required the sample to be weighted toward Western and agritourist samples, and thus, not entirely representative of the U.S. population. Beyond this limitation, there are a few areas for recommended future research. First, future studies should continue to explore the heterogeneity and relationships between the characteristics of one's surrounding area of origin and their choice behaviors, particularly the differences between rural and urban residents' choice behavior. Secondly, as mentioned throughout this dissertation, agritourism is an extremely heterogeneous industry with many unique experiences and this study has only explored a small subset of agritourism qualities. Finally, this research only considers one sector of the broader recreation market and future studies should explore the how consumer heterogeneity across space may differ for other markets of importance.

CHAPTER FIVE: CONCLUSIONS

Agritourism is an interesting sector to examine because of its connections to a variety of farm viability, community and economic development drivers. The focus of this research was identifying and analyzing factors that may strengthen linkages between rural and urban households and economies through travel behavior and expenditures. Such linkages are of interest because of the potential for diversification of farm enterprise and community economic returns based on the interdependencies of tourists' interest in farms with natural resources, rural infrastructure and other locational factors. Thus, the economic analysis, findings and discussion focus on a whole suite of regional-, household- and community-based drivers. More specifically, I identify subtle details on the heterogeneity of traveler/consumer behavior to signal a variety of opportunities for different types of farms and communities.

Agricultural diversification strategies are receiving increased attention by policymakers and rural development practitioners, and agritourism is one of the options of interest because of the varied locations and market conditions in which it can thrive. However, a particular challenge in the development of the agritourism sector, and providing technical assistance to support such efforts, has been developing a set of best practices that recognize the vast degree of variation that exists across consumers, agricultural businesses, and regions. This dissertation attempts to disentangle some of these place-based complexities by taking a three-pronged approach to analyze the sector's locational incidence, current traveler behavior, and market potential to attract more visitors given the stated preferences of US travelers.

In my first essay I consider the supply side factors that may influence agritourism with a particular focus on locational factors that influence the economic activity reported by agritourism businesses. Second, I explore three sources of heterogeneity in the industry using a revealed preference traveler demand model for agritourism. Finally, the third essay explores the market potential of US travelers through a stated preference choice model, and more specifically, identify the relative marginal values consumers place on various agritourism enterprise attributes across market segments and location.

Beyond the interesting empirical challenges faced when studying a relatively under-defined and heterogeneous product, it is important to connect these findings to the primary motivation for this dissertation is agritourism's potential to increase farm and ranch revenues, local employment, and spur regional growth. Agritourism's potential to act as a regional development tool is because, like tourism, it behaves as an "export sector" since it brings dollars to a community from outside sources. An increasing number of agricultural businesses and communities in the US, and particularly in Europe, have recognized potential producer and community benefits and begun initiatives to develop and support agritourism enterprises. Yet, the sector is still relatively understudied. The findings presented here show there is both potential for a variety of types of producers to adopt such enterprises, and that the stated and revealed value travelers place on such enterprises warrant development of a diverse range of enterprises in this sector.

Global agricultural commodity markets have become increasingly competitive and cyclical over the past several decades, adding financial pressure on small- and medium-sized farms and ranches that cannot rely on the economies of scale to compete with global prices. However, renewed interest in agriculture and foods may represent an opportunity. By leveraging a farm or ranch's locational and community assets, principal operators can tap into the public's growing interest in food systems and agriculture's heritage, to generate a new revenue stream from nearby visitors or out-of-area travelers.

In addition to this research and its empirical findings being pertinent to agricultural operators, industry stakeholders, and technical support organizations, it represents novel contributions to the regional and applied economics literature in terms of the unique methods applied. For example, the multi-destination effect in the Travel Cost Method has been debated in the recreational demand literature for decades, and while multiple studies have explored ways to work through the issue, I circumvent the specification issue by directly addressing the MD effect and capturing detailed data in my survey instrument. In general, most studies on recreation either focus on valuing one, or a small set, of sites, and subsequently, fail to sufficiently account for industry heterogeneity. Due to the rich datasets and flexible empirical models incorporated in this project, I have been able to separate out unique yet interrelated market forces in the agritourism sector, moving the field a step closer at developing a set of best practices for assessing travelers that can be customized for a diverse set of agritourism enterprises. And, this could be extended to all types of tourism when multiple destinations act as draws for travelers.

As economies around the world continue to expand, both intensively and extensively, spatial dimensions are of increasing importance to understand the dynamics of interrelated markets and general human behavior. It is for this same reason that many economists, such as Paul Krugman in his *New Economic Geography, Now Middle Aged* (2011), have called for a marriage in methods between economics and geography. Given my interest in geographic factors and the narrative one can illustrate with maps, I integrate spatial elements into each of my essays. For example, in chapter two place-based factors serve as primary variables of interest, and in chapter three, I develop a method for defining agritourism zones across the US to cluster regional agritourism market indicators. In addition, Chapter 4 presents a geostatistical interpolation method to predict WTP values for various agritourism attributes across the contiguous US. The call for greater consideration of spatial relationships in economics is particularly germane to agritourism as well as the broader tourism and outdoor recreation industry. Each of these sectors relies on the mobility and demand of local and distant visitors/travelers. And subsequently, operators can better plan and perform with an understanding of the degree to which the availability, awareness, and access to locational and community attributes will attract consumers.

In chapter two, I tested three different theories of comparative advantage to explore the effects that locational, production, and community based factors have on the decision to adopt an agritourism enterprise and its financial success. I developed what I have termed a unilateral spatial interactions model using principles developed and assessed by the seminal gravity model to describe how tourists are pulled toward individual agritourism sites across the contiguous U.S. Applying farm-level Census of Agriculture data to a two-stage Heckman model, I found that the Heckscher-Ohlin theory of comparative advantage best described the factors stimulating agritourism revenues. A particularly interesting finding in this essay, that has yet to be explored in the greater agritourism literature, is the evidence of positive externalities from agglomeration of enterprises. Translating these findings into outreach material for agritourism industry and broader community and economic development stakeholders will allow them to identify the potential region-wide draw for regions interested in developing an agritourism industry based on their natural endowments, efficiencies, and potential for cluster economies that will require them to collaborate with other tourist sites in their region.

Chapter three not only investigates heterogeneity in the agritourism industry, but also contributes to the greater research cohort refining travel cost methods by exploring and identifying a bias which has challenged the literature for decades. By adopting a flexible functional form for the travel cost model, I am able to disentangle some of the Western agritourism heterogeneity across regions, sites, and travelers. Not only does this approach allow for more precise CS estimates, but it also allows me to test for MD bias. The MD bias arises when a multi-destination visitor's travel cost calculation includes a portion of the distance already traveled that might be considered a sunk cost, instead of only measuring the incremental cost of travel.

Unlike previous methods, I address this problem by including questions in the survey instrument that directly illicit respondents' incremental travel cost information. I find that if the traveler is a MD visitor, using past methods would inflate their CS estimates by 244% on average. In addition, when the bias presented by this measurement error is not corrected for, PP CS estimates are no longer unbiased and are depressed on average by 10%. Depending on the nature of the recreational activity, the location of the activity, and the share of MD travelers, this can lead to a positive or negative bias in the overall estimated consumer benefit of the activity in question. Given the results of this chapter, future travel cost studies that do not have the benefit of using primary data should consider the implications laid out in this chapter and determine if their CS estimates may represent upper or lower bounds of the true CS values based on sample characteristics.

Chapter four applies a random utility model to a choice experiment on an overnight "ranch-stay" type agritourism trip. By using both a latent class logit model as well as a spatial interpolation technique called Empirical Bayesian Kriging (EBK), I explore the heterogeneity in agritourism preferences across market segments as well as across space. Four classes, defined as the Comfort Tourists, Adventurers, Road Trippers, and Urbanites, are identified, and their preferences are found to be significantly driven by the respondent's home region, and their surrounding agritourism activity, income per capita, age, and natural amenities.

While identifying the differences in preferences across market segments, it shows opportunities for a variety of agricultural operations essentially identifying geographical markets with relatively higher WTP values for various activities might also enable stakeholders to become more effective at target marketing to regions with the greatest potential interest. Class weighted WTP estimates for some place-based attributes, such as the proximity to NPS lands and urban areas, do not vary significantly across space. However, the WTP for other enterprise attributes, such as offering tent lodging and horseback riding, vary significantly across the contiguous US. Based on a thorough scan of the literature, this may be the first study to map the WTP for attributes of a recreational activity. Mapping the WTP for other recreational activities, which may vary by place and whose participants' WTP values are likely influenced by their experience with similar or contrasting place-based factors is warranted in future research.

Despite its potential as a farm diversification and rural development tool, agritourism is not a magic bullet, but may be worth exploring for farms in a variety of locations, depending on their characteristics and interests in various enterprises. The findings in the three essays of this dissertation indicate that small- and medium-sized farms and ranches near population centers in regions with attractive natural endowments are the most likely to benefit from agritourism. Yet, these urban adjacent communities already tend to fair better than those in more remote locations due to their greater access to urban buying dollars, capital, and employment opportunities. More research should be conducted on whether agritourism should be among the supported economic opportunities for what Cromartie, Von Reichert, & Arthun (2015) aptly names "geographically disadvantaged" counties, counties that are both remote and lack scenic amenities.

Even if an agritourism farm or ranch is not near a population center, chapters two and three show how it can still benefit from drawing multi-destination travelers passing through their area, particularly if situated near a well-traveled transportation corridor or adjacent to key outdoor recreation assets (National parks, forests and/or seashores). Multi-destination travelers may be less likely to travel long distances off their primary route, however, there is still potential for primary destination agritourism operations to achieve financial success if the operators cater their enterprise development and marketing efforts toward traveler segments and geographic markets with higher WTP values for the activities they offer.

Agritourism has many linkages with other important industries for rural economies that should be further explored, including agritourism's relationship with National Parks, outdoor recreation, and entrepreneurship. Along with Van Sandt et al. (2018), chapter two provides evidence for potential partnerships between the National Parks Service and agritourism organizations. The higher agritourism revenues generated by agritourism operations near National Parks is most likely due to multi-destination tourists bundling complementary recreation experiences on their trips, signaling opportunities for joint-marketing efforts between agritourism establishments and other outdoor recreation sites. Among other analyses, case studies should be conducted on particular regions of interest to identify the potential spillover benefits, opportunities, and barriers between agritourism operations and outdoor recreation sites.

Finally, more research should be conducted on agritourism's relationship with rural entrepreneurship. In the U.S., agritourism and outdoor recreation, have connections to larger rural policy issues including land use and conservation, emerging public interest in reconnecting with their food system, and leveraging the diverse heritage and histories of rural communities across America to create entrepreneurial and economic opportunities. The apparent role of community-based assets explored in chapter two and the spatial distribution of WTP for agritourism activities in chapter four suggest a need for case studies and surveys that focus on how either entrepreneurs or collaboratives of operators can strategically leverage their tourism development activities to achieve sustainable regional economic growth.

REFERENCES

AAA (2014). Your Driving Costs: How much are you really paying to drive? 1000 AAA Drive, Heathrow, FL. http://exchange.aaa.com/wp-content/uploads/2014/05/Your-Driving-Costs-2014.pdf

Adamowicz, W.L., Fletcher, J.J. & Graham-Tomasi, T. (1989). 'Functional form and the statistical properties of welfare measures', *American Journal of Agricultural Economics*, vol. 71, pp. 414–421.

Arrow, K. J. (1962). The economic implications of learning by doing. *The review of economic studies*, *29*(3), 155-173.

Arroyo, C. G., Barbieri, C., & Rich, S. R. (2013). Defining agritourism: A comparative study of stakeholders' perceptions in Missouri and North Carolina. *Tourism Management*, *37*, 39-47.

Bagi, F.S. & R. Reeder (2012a). Farm Activities Associated With Rural Development Initiatives*,* ERR-134, U.S. Department of Agriculture, Economic Research Service, May. 48*pp*.

Bagi, F. S., & R. Reeder (2012b). Factors affecting farmer participation in agritourism. *Agricultural and Resource Economics Review*, *41*(2), 189.

Barbieri, C. (2013). Assessing the sustainability of agritourism in the US: a comparison between agritourism and other farm entrepreneurial ventures. *Journal of Sustainable Tourism*, *21*(2), 252-270.

Barbieri, C., & E. Mahoney (2009). Why is diversification an attractive farm adjustment strategy? Insights from Texas farmers and ranchers. *Journal of rural studies*, *25*(1), 58-66.

Bateman, I.J., B.H. Day, A.P. Jones, & S. Jude (2009). "Reducing gain–loss asymmetry: a virtual reality choice experiment valuing land use change." *Journal of Environmental Economics and Management* 58 (1) 106-118.

Bauman, A., Thilmany, D., Jablonski, B.B.R., & Shideler, D. (2015). An Overview of Emerging Business Models in the Local Foods Landscape. Smart Marketing.

Benson, C., P. Watson, G. Taylor, P. Cook, & S. Hollenhorst (2013). "Who visits a national park and what do they get out of it?: A joint visitor cluster analysis and travel cost model for Yellowstone National Park." *Environmental management* 52 (4), 917-928.

Bhat, C. R. (1997). An endogenous segmentation mode choice model with an application to intercity travel. *Transportation science*, *31*(1), 34-48.

Bishop, R. C., & Heberlein, T. A. (1979). Measuring values of extramarket goods: Are indirect measures biased?. *American journal of agricultural economics*, *61*(5), 926-930.

Bockstael, N.I. Strand & M. Hanemann, (1987). ’’Timeand The Recreation Demand Model,’’ American *Journal of Agricultural Economics 69(2):293–302.*

Brouwer, R., Martin-Ortega, J., & Berbel, J. (2010). Spatial preference heterogeneity: a choice experiment. *Land Economics*, *86*(3), 552-568.

Brown, J. P., S. J. Goetz, M. C. Ahearn, and C. Liang (2014). Linkages between community-focused agriculture, farm sales, and regional growth. *Economic Development Quarterly* 28 (1): 5-16.

Busby, Graham, & Samantha Rendle. "The transition from tourism on farms to farm tourism." *Tourism management* 21, no. 6 (2000): 635-642.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge university press.

Campbell, D. (2007). Willingness to Pay for Rural Landscape Improvements: Combining Mixed Logit and Random-Effects Models. *Journal of agricultural economics*, *58*(3), 467-483.

Carpio, C., Wohlgenant, M., and Boonsaeng, T. (2008), 'The demand for agritourism in the United States', *Journal of Agricultural and Resource Economics*, Vol 33, No 2, pp 254–269.

Carreras, C., and Verdaguer, C. (1995). Mega-events: Local strategies and global tourist attractions. *European tourism: Regions, spaces and restructuring*, 193-205.

Caussade, S., de Dios Ortúzar, J., Rizzi, L. I., & Hensher, D. A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation research part B: Methodological*, *39*(7), 621-640.

Chatterji, A., Glaeser, E., & Kerr, W. (2014). Clusters of entrepreneurship and innovation. *Innovation Policy and the Economy*, *14*(1), 129-166.

Che, D., Veeck, A., & Veeck, G. (2005). Sustaining production and strengthening the agritourism product: Linkages among Michigan agritourism destinations. *Agriculture and Human values*, *22*(2), 225-234.

Cromartie, J., Von Reichert, C., & Arthun, R. (2015). *Factors affecting former residents' returning to rural communities* (No. 206008). United States Department of Agriculture, Economic Research Service.

Das, B. R., & Rainey, D. V. (2010). Agritourism in the Arkansas Delta Byways: Assessing the economic impacts. *International Journal of Tourism Research*, *12*(3), 265-280.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic Review*, *97*(1), 197-221.

Fourie, J., & Santana-Gallego, M. (2013). Ethnic reunion and cultural affinity. *Tourism Management*, *36*, 411-420.

Frosch, B. J., Anderson, D. P., & Outlaw, J. L. (2008, February). Economic impact of deer breeding operations in Texas. In *Southern Agricultural Economics Association Annual Meeting* (pp. 2-6).

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of political economy*, *100*(6), 1126-1152.

Glaeser, E., S. Kerr, & W. Kerr (2012). "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines." NBER Working Paper no. 18333. Cambridge, MA: National Bureau of Economic Research.

Gray, H. P. (1989). Services and comparative advantage theory. *Services in world economic growth*, 65-103.

Grogger, J. T., & Carson, R. T. (1991). Models for truncated counts. *Journal of applied econometrics*, *6*(3), 225-238.

Hanley, N., Wright, R. E., & Koop, G. (2002). Modelling recreation demand using choice experiments: climbing in Scotland. *Environmental and resource Economics*, *22*(3), 449-466.

Haspel, A. E., & Johnson, F. R. (1982). Multiple destination trip bias in recreation benefit estimation. *Land economics*, *58*(3), 364-372.

Hensher, D. A., & Greene, W. H. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. *Empirical economics*, *39*(2), 413-426.

Hidrue, M. K., Parsons, G. R., Kempton, W., & Gardner, M. P. (2011). Willingness to pay for electric vehicles and their attributes. *Resource and Energy Economics*, *33*(3), 686-705.

Hill, R., Loomis, J., Thilmany, D., & Sullins, M. (2014). Economic values of agritourism to visitors: a multi-destination hurdle travel cost model of demand. *Tourism Economics*, *20*(5), 1047-1065.

Hoppe, R & J. MacDonald (2013). Updating the ERS Farm Typology, EIB-110, U.S. Department of Agriculture, Economic Research Service.

Hynes, S., Hanley, N., & Scarpa, R. (2008). Effects on welfare measures of alternative means of accounting for preference heterogeneity in recreational demand models. *American journal of agricultural economics*, *90*(4), 1011-1027.

Jacobs, J. (1969). The Economy of Cities. Random House, New York.

Juutinen, A., Mitani, Y., Mäntymaa, E., Shoji, Y., Siikamäki, P., & Svento, R. (2011). Combining ecological and recreational aspects in national park management: A choice experiment application. *Ecological economics*, *70*(6), 1231-1239.

Keum, K. (2010). Tourism flows and trade theory: a panel data analysis with the gravity model. *The Annals of Regional Science*, *44*(3), 541-557.

Kimura, F., & Lee, H. H. (2006). The gravity equation in international trade in services. *Review of world economics*, *142*(1), 92-121.

Klein, A., and Crafts, N. (2012). Making sense of the manufacturing belt: determinants of US industrial location, 1880–1920. *Journal of Economic Geography*, *12*(4), 775-807.

Krivoruchko, K. (2012). Empirical bayesian kriging. *ArcUser Fall*, 6-10.

Krugman, Paul. "The new economic geography, now middle-aged." *Regional Studies* 45, no. 1 (2011): 1-7.

Lasandahasi T. de Silva & Sarath S. Kodithuwakku (2011). Pluriactivity, Entrepreneurship, and Socio-Economic Success in Farming Households. In Alsos, G. A., Carter, S., and Ljunggren,

E. (Eds.), *The handbook of research on entrepreneurship in agriculture and rural development (pp. 38-53)*. Edward Elgar Publishing.

Little, R. J., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John Wiley & Sons.

Loomis, J., & Walsh R. (1997), *Recreation Economic Decisions: Comparing Benefits and Costs*, 2nd edn, Venture Press, State College, PA.

Loomis, J., Yorizane, S., & Larson D. (2000), 'Testing significance of multi-destination and multipurpose trip effects in a travel cost method demand model for whale watching trips', *Agricultural and Resource Economics Review*, Vol 29, No 2, pp 183–191.

Loomis, J. (2006). A comparison of the effect of multiple destination trips on recreation benefits as estimated by travel cost and contingent valuation methods. *Journal of Leisure Research*, *38*(1), 46-60.

Low, S. A., Adalja, A., Beaulieu, E., Key, N., Martinez, S., Melton, A., and Jablonski, B.B.R. (2015). Trends in US local and regional food systems: A report to Congress.

Low, S. A., J. Henderson, and S. Weiler (2005). Gauging a Region's Entrepreneurial Potential. *Federal Reserve Bank of Kansas City Economic Review* 90(3), 61-89.

Marshall, A. (1890). The Principles of Economics. Macmillan, London.

Martínez-Espiñeira, R., & Amoako-Tuffour, J. (2008). Recreation demand analysis under truncation, overdispersion, and endogenous stratification: An application to Gros Morne National Park. *Journal of environmental management*, *88*(4), 1320-1332.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, *3*(4), 303-328.

McGehee, N. G., & Kim, K. (2004). Motivation for agri-tourism entrepreneurship. *Journal of travel research*, *43*(2), 161-170.

McGranahan, D. A., Wojan, T. R., & Lambert, D. M. (2010). The rural growth trifecta: outdoor amenities, creative class and entrepreneurial context. *Journal of Economic Geography*, *11*(3), 529-557.

Mendelssohn, R.J. Hof, G. Peterson and Johnson, R. (1992). "Measuring Recreation Values with Multiple Destination *Trips.'' American Journal of Agricultural Economics 74: 926-933.*

Morley, C. L. (1994). Discrete choice analysis of the impact of tourism prices. *Journal of Travel Research*, *33*(2), 8-14.

Morley, C., Rosselló, J., & Santana-Gallego, M. (2014). Gravity models for tourism demand: theory and use. *Annals of Tourism Research*, *48*, 1-10.McGranahan citation

Nelson, F. D. (1984). Efficiency of the two-step estimator for models with endogenous sample selection. *Journal of Econometrics*, *24*(1-2), 181-196.

Nickerson, N. P., Black, R. J., & McCool, S. F. (2001). Agritourism: Motivations behind farm/ranch business diversification. *Journal of Travel research*, *40*(1), 19-26.

NPS (2014). Visitor Use Statistics. Retrieved from: https://irma.nps.gov/Stats/

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling*, *14*(4), 535-569.

OECD (2016). An Integrated Approach to Economic Development. Retrieved from http://www.oecd.org/gov/rural-urban-linkages.htm.

Pacifico, D., & Yoo, H. (2013). lclogit: a Stata command for fitting latent-class conditional logit models via the expectation-maximization algorithm. *Stata journal.*, *13*(3), 625-639.

Parsons, G. R., & Wilson, A. J. (1997). Incidental and joint consumption in recreation demand. *Agricultural and Resource Economics Review*, *26*, 1-6.

Phaneuf, D. J., & Smith, V. K. (2005). Recreation demand models. *Handbook of environmental economics*, *2*, 671-761.

Puhani, P. (2000). The Heckman correction for sample selection and its critique. *Journal of economic surveys*, *14*(1), 53-68.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, *94*(5), 1002-1037.

Rosenthal, D. H. (1987). The necessity for substitute prices in recreation demand analyses. *American Journal of Agricultural Economics*, 828-837.

Shen, J. (2009). Latent class model or mixed logit model? A comparison by transport mode choice data. *Applied Economics*, *41*(22), 2915-2924.

Schilling, B. J., Sullivan, K. P., & Komar, S. J. (2016). Examining the economic benefits of agritourism: The case of New Jersey. *Journal of Agriculture, Food Systems, and Community Development*, *3*(1), 199-214.

Stahl, K. (1982). Differentiated Products, Consumer Search, and Locational Oligopoly. *Journal of Industrial Economics*, 97-113.

Skuras, D., E. Dimara, & A. Petrou (2006). "Rural tourism and visitors' expenditures for local food products." *Regional Studies* 40, no. 7: 769-779.

Smith, V.K., & Kopp R. (1980), "The Spatial Limits of the Travel Cost Recreational Demand Model." *Land Economics* 56:64-72.

Tew, C., & Barbieri, C. (2012). The perceived benefits of agritourism: The provider's perspective. *Tourism Management*, *33*(1), 215-224.

Thilmany, D., Sullins, M., & Ansteth, A. (2007). The 2006 economic contribution of agritourism to Colorado: Estimates from a survey of Colorado tourists. *Economic development report (Colorado State University. Dept. of Agricultural and Resource Economics); EDR 07-24*.

Train, K. E. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, *1*(1), 40-69.

Urtasun, A., & Gutiérrez, I. (2006). Tourism agglomeration and its impact on social welfare: An empirical approach to the Spanish case. *Tourism Management*, *27*(5), 901-912.

U.S. Dept. of Agriculture (2014a). Ag census: desktop data query tool [Data file]. Retrieved from: http://www.agcensus.usda.gov/Publications/2012/

USDA (2014b). Rural Atlas Data [Data file]. Retrieved from: http://www.ers.usda.gov/dataproducts/ atlas-of-rural-and-small-town-america.aspx

USDA, Economic Research Service (2018). *Rural America at a Glance: 2017 Edition*. Retrieved from https://www.ers.usda.gov

U.S. Dept. of Transportation, Office of Aviation Analysis (2014). Domestic Airline Fares Consumer Report. Report accessed May 2015 at: https://www.transportation.gov/sites/dot.gov/files/docs/Consumer%20Airfare%20Report%20Dec%202013%20Rev.pdf

van der Panne, G. (2004). Agglomeration externalities: Marshall versus jacobs. *Journal of Evolutionary Economics*, *14*(5), 593-604.

Van Sandt, A., & Thilmany, D. (2016). Diversification through Agritourism in a Changing US Farmscape. *Western Economics Forum*, *15*(1), 52.

Van Sandt, A., Low, S. A., & Thilmany, D. (2018). Exploring Regional Patterns of Agritourism in the US: What's Driving Clusters of Enterprises?. *Agricultural and Resource Economics Review*, 1-18.

Veeck, G., Che, D., & Veeck, A. (2006). America's Changing Farmscape: A Study of Agricultural Tourism in Michigan*. *The Professional Geographer*, 58(3), 235-248.

Verheul, I., Thurik, R., Hessels, J., and van der Zwan, P. (2010). Factors influencing the entrepreneurial engagement of opportunity and necessity entrepreneurs. *EIM Research Reports h*, *201011*, 1-24.

Weiler, Stephan (2017). Six Charts that Illustrate the Divide Between Rural and Urban America. Colorado State University, Source. Report accessed June, 2017 at: http://source.colostate.edu/six-charts-illustrate-divide-rural-urban-america/

Wen, C. H., & Lai, S. C. (2010). Latent class models of international air carrier choice. *Transportation Research Part E: Logistics and Transportation Review*, *46*(2), 211-221.

Wilson, J., Thilmany, D., & Sullins, M. (2006). Agritourism: A potential economic driver in the rural west. *Economic development report (Colorado State University. Dept. of Agricultural and Resource Economics); EDR 06-01*.

Zhang, J., & Jensen, C. (2007). Comparative advantage: explaining tourism flows. *Annals of tourism research*, *34*(1), 223-243.

APPENDIX

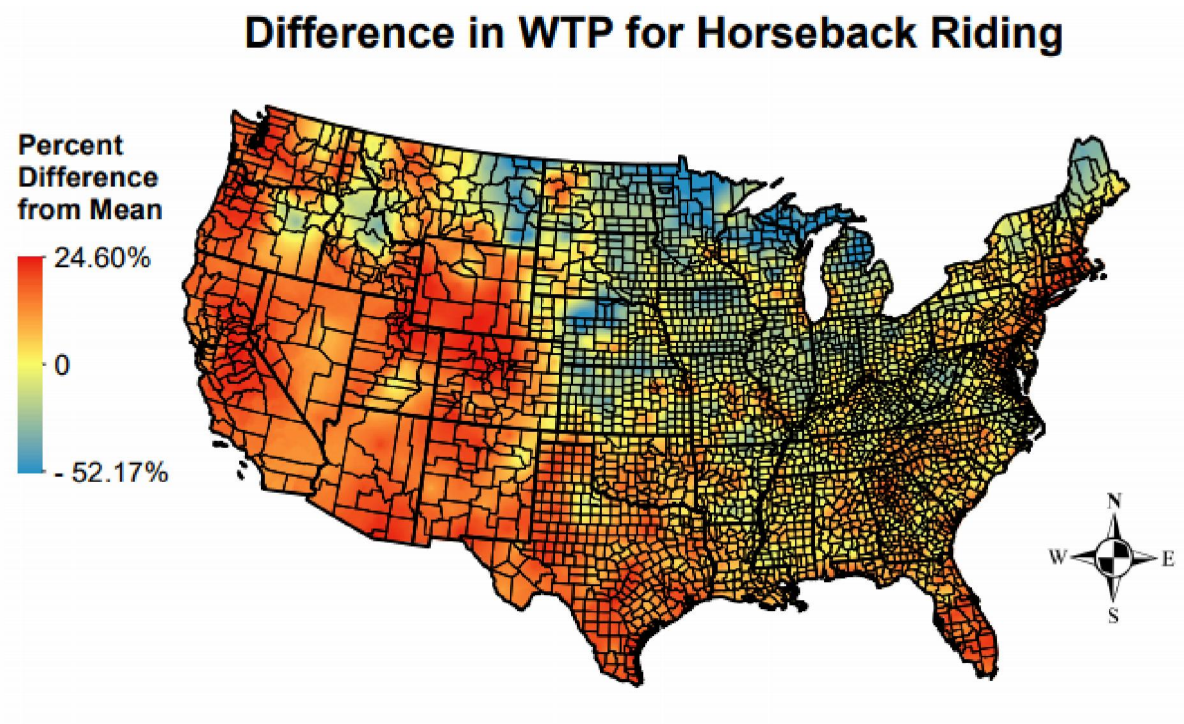


Figure A.1 Difference in Willingness to Pay for Horseback Riding

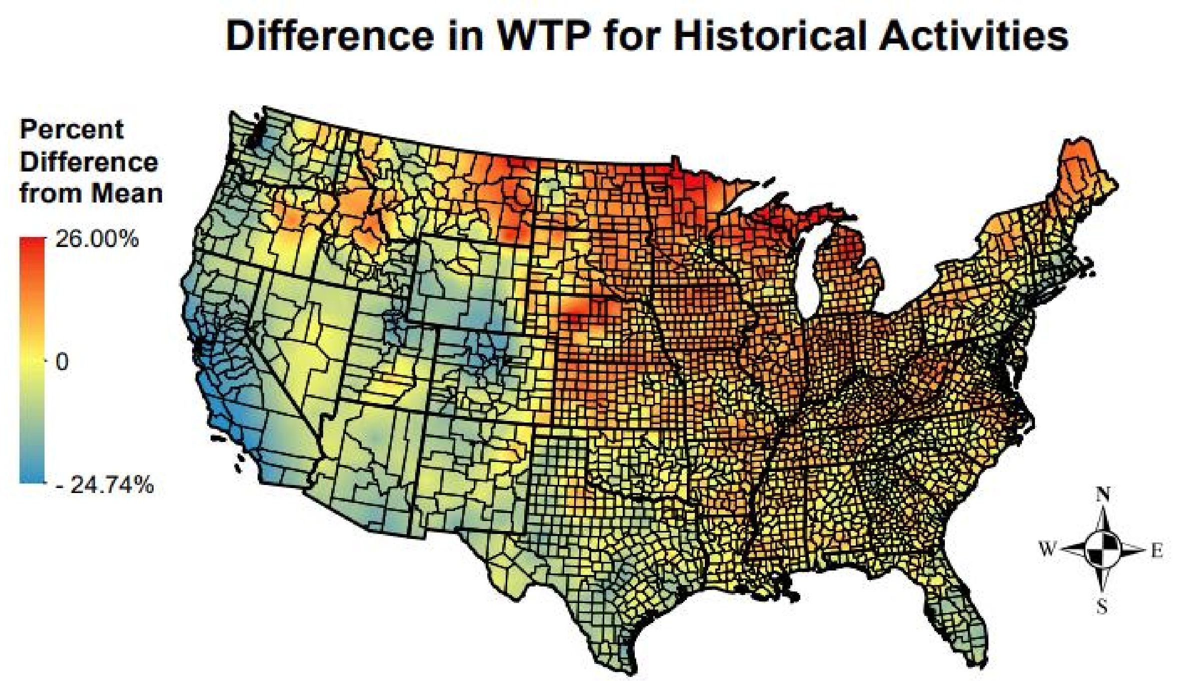


Figure A.2 Difference in Willingness to Pay for Historical Activities

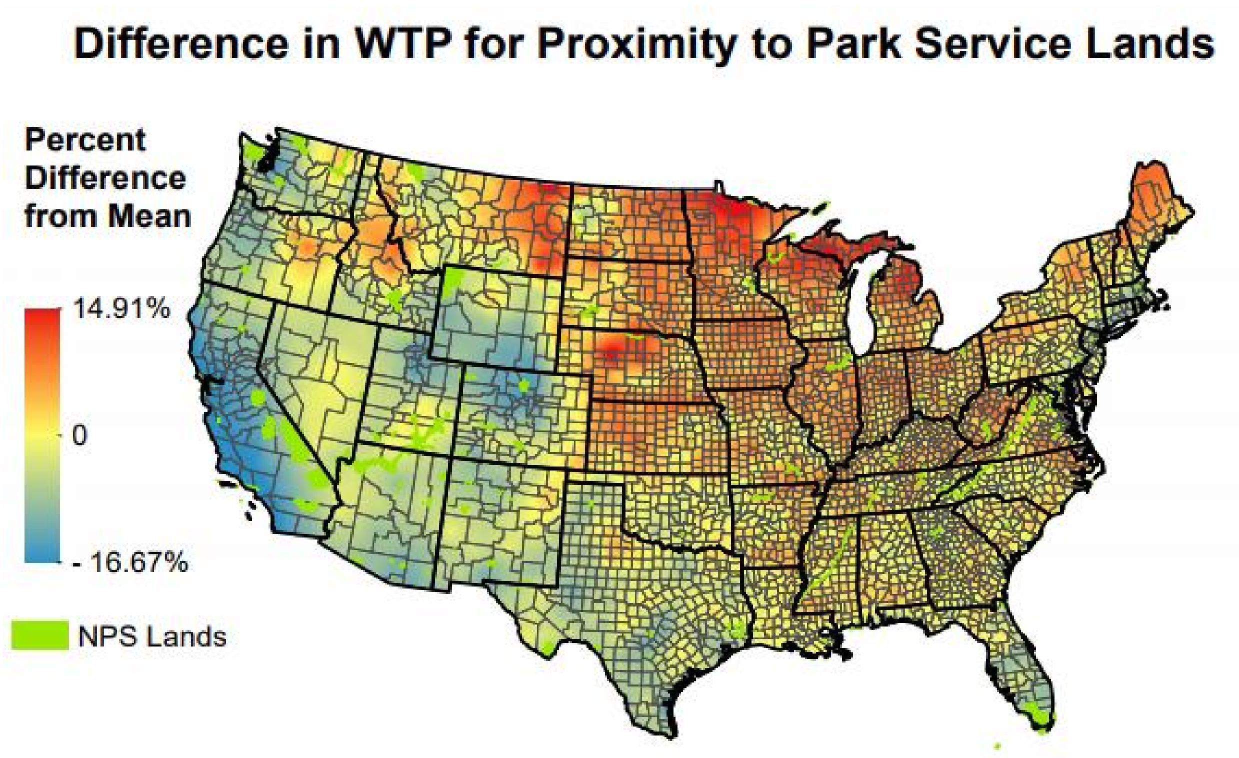


Figure A.3 Difference in Willingness to Pay for Proximity to Park Service Lands

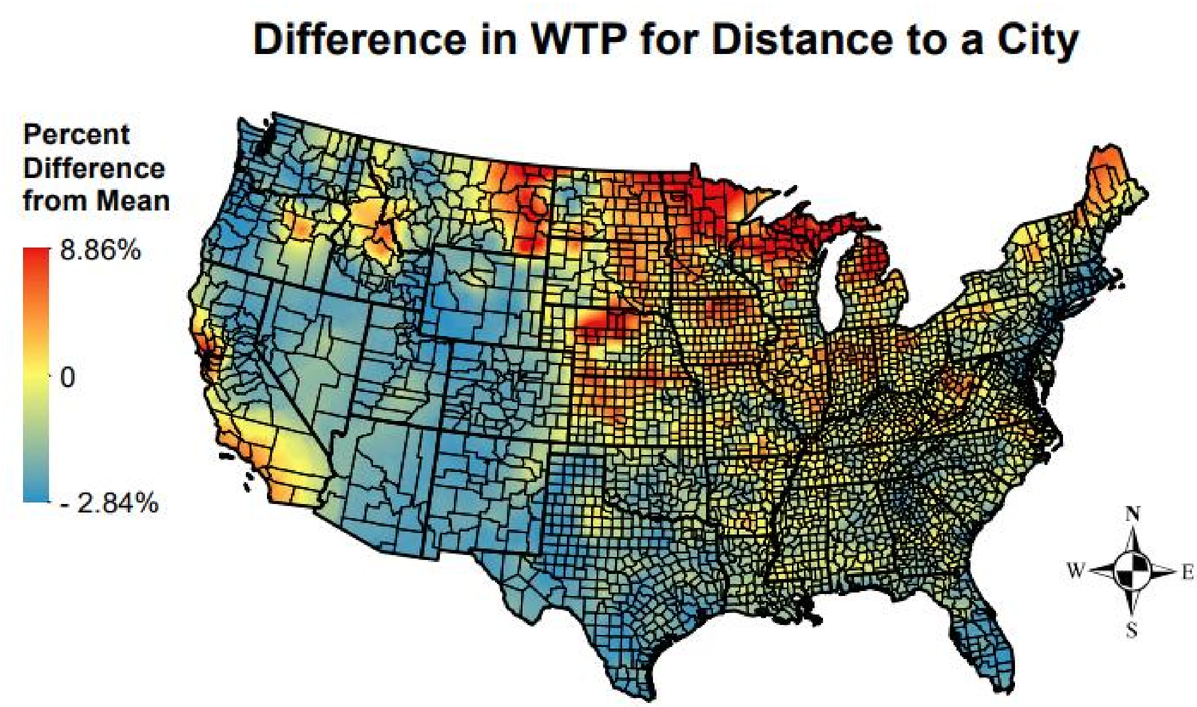


Figure A.4 Difference in Willingness to Pay for Distance to a City

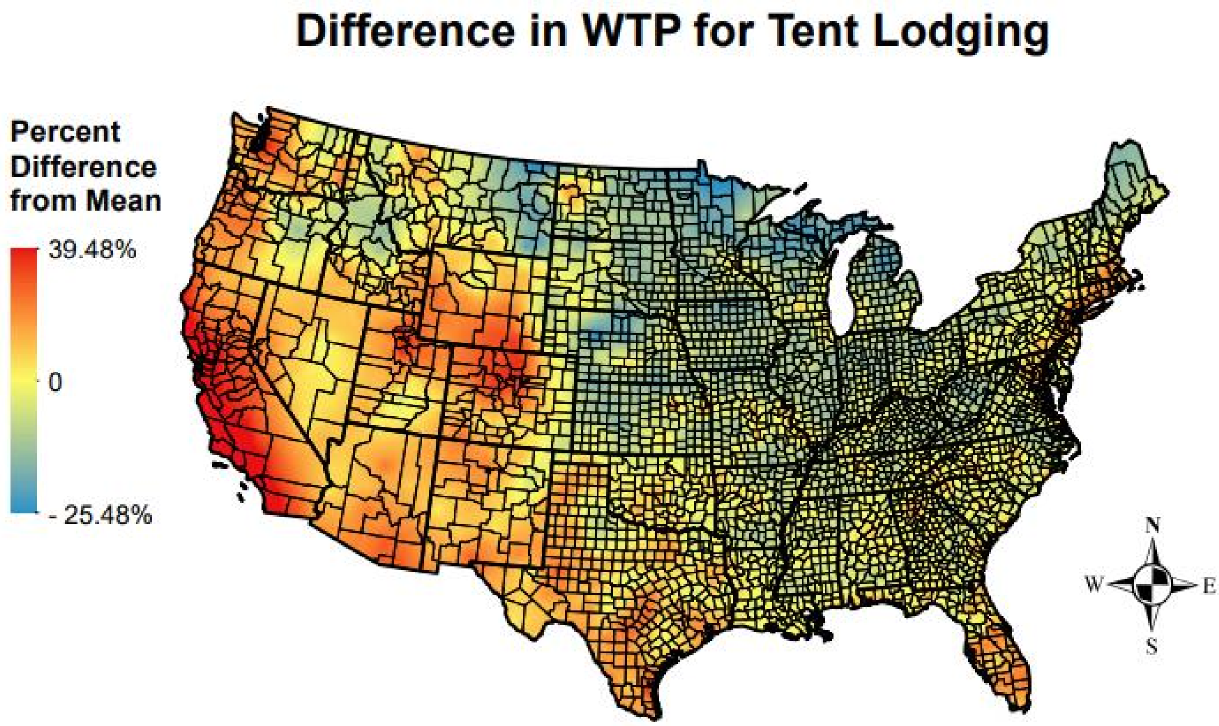


Figure A.5 Difference in Willingness to Pay for Tent Lodging

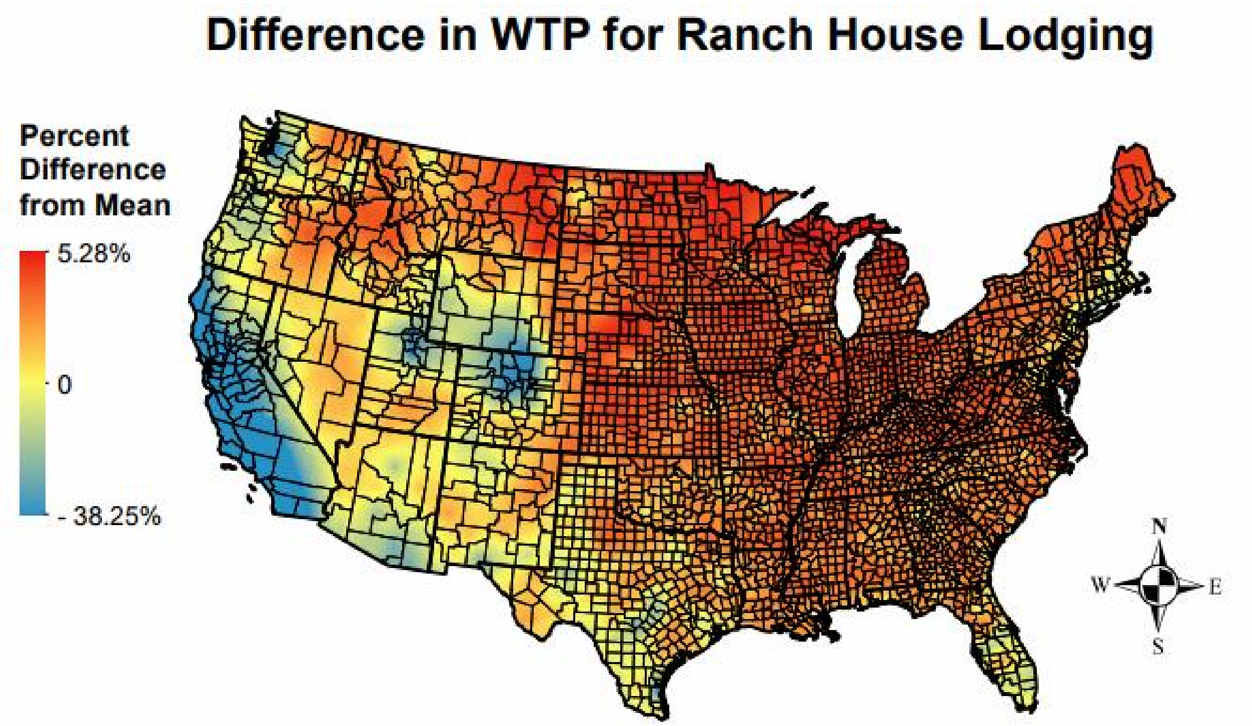


Figure A.6 Difference in Willingness to Pay for Ranch House Lodging

Technical Appendix

The main statistical tool applied in the kriging method is the semivariogram which plots the pair-wise differences (semivariances) separated by some distance, h. Plotting the pair-wise semivariances over distance in a semivariogram shows how these dissimilarities increase as distance increases, as shown in the concave shape of the semivariogram (figure A.10). There are three main components, or parameters of the semivariogram: the nugget, range, and sill. The nugget is the y-intercept on the graph and should theoretically always be zero since at a distance of zero there should be zero autocorrelation, however, in empirical applications the nugget frequently takes positive values. At some point, increasing the distance past a certain threshold, called the range, does not increase the semivariance, thus observations beyond the range are said to be uncorrelated. Finally, the sill can be thought of as the ceiling of autocorrelation associated with the range. In other words, the sill is the average semivariance that occurs past the range once the semivariogram levels out. Kriging makes use of these parameters to provide measures of dissimilarity and develop weights that can then be used to predict un-sampled points.

(A.1) $$\hat{\varphi}(h) = \frac{1}{2} * \frac{1}{N(h)} \sum_{i=1}^{N(h)} (z_i - z_{i+h})^2$$

The semivariance is defined in equation A.1, where, h is a pre-specified distance; $\hat{\varphi}(h)$ is the semivariance given h; $N(h)$ is the number of observations in a neighborhood defined by a radius of h; and z_i is the value of a variable at a known point. (The reader may recognize equation A.1 is very similar to the equation for mean squared error). While I do not make use of the ordinary kriging estimator, which I explain the reasoning for shortly, it may be help facilitate comprehension if the process of estimating weights in this method is briefly explained. In its simplest form, once the semivariances are calculated for each pair of points, we have a measure of the dissimilarities of the points that we will attempt to minimize using a Lagrangian optimization where the constraint is that the weights sum to one. By taking the first order conditions with respect to the Lagrangian multiplier and the weights and setting them equal to zero, we arrive at a result where the semivariances between points is minimized and the weights sum to one. When we arrive at this solution, the kriging estimates are said to be the best linear unbiased estimates among all other spatial interpolation predictors.

Issues arise, however, when the distribution of the data is unknown and when there is no constant global mean (non-stationarity). Intuitively, non-stationarity implies that spatial autocorrelation behaves differently depending on the neighborhood being observed. This is likely considering the spatial extent of the sample includes all 48 states in the contiguous U.S. Geographical barriers such as large mountain ranges and deserts are likely to disturb the continuity of preferences over space, thus leading to different patterns of spatial correlation across space. When stationarity in the data breaks down and the assumed distribution of the data is incorrect, ordinary kriging is no longer BLUE, motivating us to pursue other research methods. One such method is empirical Bayesian kriging (EBK).

In EBK the sample is divided into subsets, or neighborhoods so that the kriging weights in different neighborhoods are not pulling the data toward one global mean, but instead represent the nature of the semivariogram that may be unique to that neighborhood. Some overlapping of these neighborhoods is allowed to provide some continuity over space, but again, the main goal of this step is to attempt to correct for mild degrees of non-stationarity. Following estimation of the semivariograms for each neighborhood using the dataset, new points of data are generated within each of the neighborhoods and another semivariogram is created based on this simulated data. This step is iterated many times resulting in an estimated distribution for each of the semivariograms. Given the nature of the simulated data, the kriging parameters (range, sill, and nugget) for EBK must be calculated using mixed maximum likelihood estimation because of the within-subject and between-subject variation. The within-subject variation is a product of the simulated points while the between-subject variation refers to the relationships between simulated clusters of points across the spatial extent of the neighborhood. Once a semivariogram is estimated for a neighborhood, weights can be calculated to estimate un-sampled points. Again, the benefits of this method are that a distribution need not be assumed, and non-stationarity in the data that may arise from the geographical scope of the data set (the contiguous U.S.) or large geographical boundaries (e.g. mountain ranges, deserts, and great lakes) can be somewhat controlled for.

Table A.1 Empirical Bayesian Kriging Summary Statistics

Prediction Error Statistics	Historical Excursion	Horseback Riding	Proximity to City	Proximity to Park	Tent	Ranch House
Mean	-0.0004	0.0008	-0.0001	0.0002	0.0004	-0.0002
Root Mean Square	0.0491	0.0871	0.0148	0.0289	0.0533	0.0288
Mean Standardized	-0.0029	0.0033	-0.0007	0.0030	0.0036	-0.0055
Root Mean Square Standardized	0.9925	0.9881	0.9829	0.9936	0.9958	0.9905
Avg. Standard Error	0.0491	0.0863	0.0143	0.0289	0.0531	0.0273

Table A.2 Heckman Selection Equation – Model Two
Dependent variable: 1 if Agritourism, 0 otherwise

	Variable	Marginal Effects
Farm level	Intercept	-0.0495***
	Ln(total value of production)	10.1873***
	Female	0.1231***
	Black	-0.0104*
	Asian	0.0045
	Hawaiian	0.0004
	American Indian	-0.0205***
	Retired	-0.3002***
	Age	59.1551***
	Days worked off farm	-2.9185***
	Crops and Livestock	
Farm level	Hay and grains	-0.3922***
	Christmas trees	0.0076***
	Maple products	-0.0032**
	Bee products	0.0196***
	Vegetables	0.0361***
	Fruit and Nuts	0.0561***
	Berries	0.0155***
	Grapes	0.0023***
	Cattle	-0.4207*
	Horses	0.2081***
	Sheep and goats	0.0757***
	Pigs	0.0346*
	Poultry	-0.1033***
	Other livestock	0.0191***
	Forest products	0.0262***
	Value added products	0.0431***
	Direct to consumer	0.0681***
	Direct to retailer	0.0244***
	Organic certified	0.0108*
	On farm packaging facility	0.0116***
	Spatial	
County level	South	0.3248***
	Midwest	-0.4152***
	Northeast	0.1113***
	Miles of Byways/100 sq. mi.	1.7221***
	(Miles of Byways/100 sq. mi.)^2	-10.6574***
	Miles of interstates/100 sq. mi.	-1.6878
	(Miles of interstates/100 sq. mi.)^2	-7.0334
	Ln(population)	-10.6513***
	Farm dependent	0.1318***

Table A.2 continued…	
Recreation dependent	0.0961***
Entrepreneurship	
Breadth	0.2395***
Patents per 1,000 people	1.4783***
Observations: 1,078,115	
* Significant at 1% level	
** Significant at 5% level	
*** Significant at 10% level	

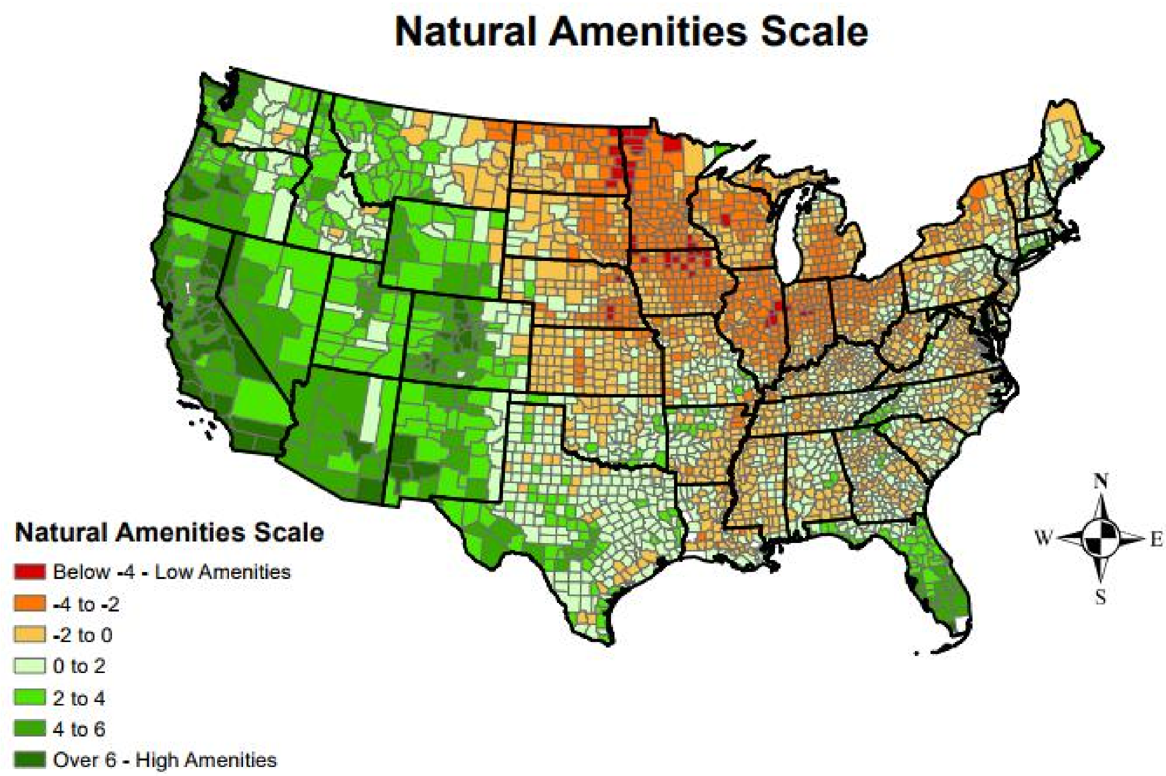
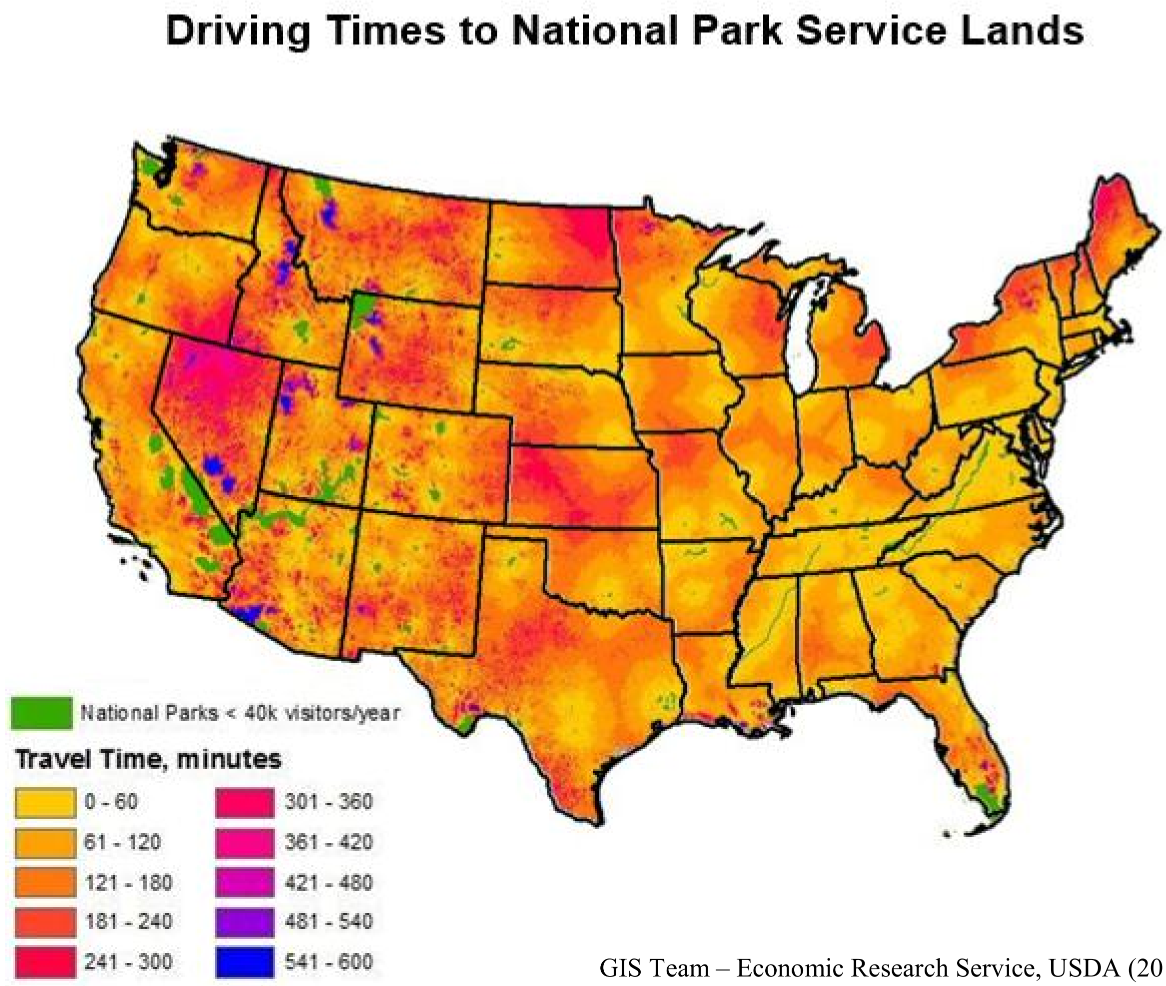


Figure A.7 McGranahan's Natural Amenities Scale



GIS Team – Economic Research Service, USDA (2015)

Figure A.8 Driving Times to National Park Service Lands

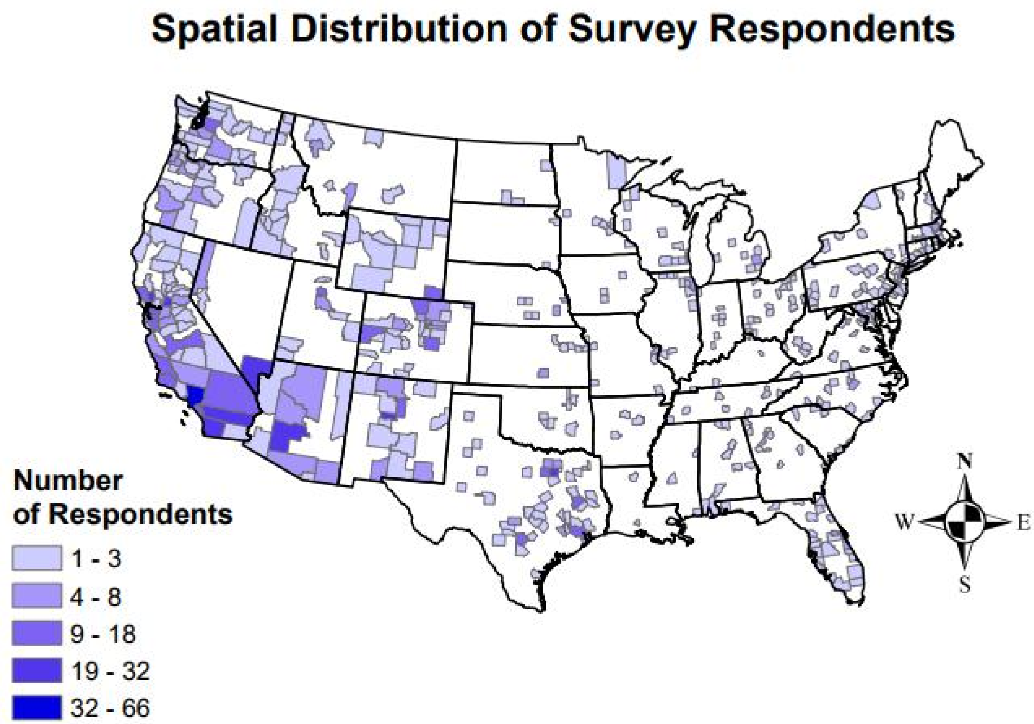


Figure A.9 Spatial Distribution of Survey Respondents

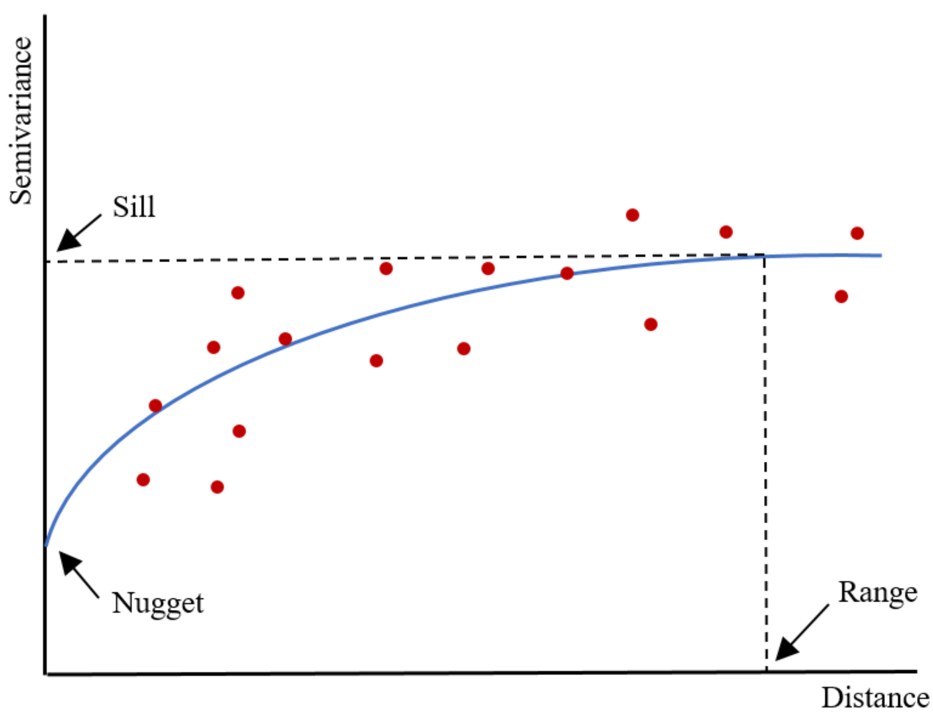


Figure A.10 Example of a Semivariogram