CONSUMER PERCEPTIONS, PREFERENCES, AND WILLINGNESS TO PAY FOR SUSTAINABLY LABELED FOOD PRODUCTS

By

Hillary M. Sackett

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food, and Resource Economics - Doctor of Philosophy

2013

ABSTRACT

CONSUMER PERCEPTIONS, PREFERENCES, AND WILLINGNESS TO PAY FOR SUSTAINABLY LABELED FOOD PRODUCTS

By

Hillary M. Sackett

"Sustainably produced" food labels have rapidly grown in popularity over the past decade. However, because there is no government agency overseeing certification of these production methods, consumers may be generally confused about the production attributes that may or may not be present in a sustainable food system. Moreover, the ability of a firm to differentiate their product depends critically on an accurate understanding of the perceptions consumers hold about the implications of a credence labeling claim. Building upon existing work evaluating other food attribute labels and the impact of consumer inferences, this work begins to address gaps in the literature regarding food products with sustainable production claims.

To achieve these goals, a comprehensive set of information was gathered from consumers, through two complementary activities: a nation-wide survey on food consumption habits including simulated shopping scenarios designed to reveal tradeoffs perceived by the consumer and a valuation field experiment designed to capture the same information in an in-person, non-hypothetical setting. The juxtaposition of the hypothetical and non-hypothetical results enhances the reliability and appropriateness of the economic analysis and marketing recommendations. Collectively, these two approaches yield a more detailed understanding of consumer behavior.

The first essay initiates the process of examining consumer inferences and valuations of food products making "sustainably produced" claims utilizing a Best-Worst scaling framework to identify what consumers believe "sustainably produced" labels mean and their preferences for each of the sustainable farming practices considered. The results of which suggest that consumers perceive farm size and local production as highly important elements of sustainable agriculture. Additionally, consumer preferences over economic attributes such as consumer food prices and financial stability of farmers exhibit high heterogeneity, indicating segmentation in the sample and potential for targeted marketing management.

The second essay analyzes data from a hypothetical choice experiment to better understand consumer purchasing behavior when faced with competing food production attributes such as "sustainable", "organic" and "local". The primary goal in this piece is to estimate preferences for "sustainably produced" food products and determine how they may be affected by varying degrees of information about sustainable agricultural systems. Additionally, the willingness to pay measurements estimated in this essay provide insight into the trade-offs perceived between current eco-labeling schemes, and the potential for differentiating "sustainably produced" products from their "organic" and "local" counterparts.

The third and final essay compares the results of the experimental auction and the hypothetical choice experiment to further examine the influence of sustainable, organic, and local production labels on food choice and to estimate consumer willingness to pay for sustainably produced apples and steak. The findings suggest a preference ordering in experimental bids, differing from that found in the hypothetical choice experiment. However, upon deeper investigation, the estimates do not provide sufficient evidence of hypothetical bias in survey responses. Furthermore, the data indicate statistically insignificant differences in willingness to pay for the sustainable, organic, and local products, implying a weak level of differentiation between labels. Overall, this work supports other experimental findings in the literature of positive price premiums on these credence attributes.

ACKNOWLEDGEMENTS

I would like to thank my committee, namely Dr. Robert Shupp, Dr. Glynn Tonsor, Dr. Phil Howard, and Dr. Robert Brent Ross for their guidance in this research. The entire committee provided feedback and additional insight that resulted in higher quality output. I am especially appreciative of Dr. Shupp's contributions to this work and for taking on the role of teacher by introducing me to the field of experimental economics, the role of mentor through his support and encouragement as my major professor in my academic program, and for giving me trust and freedom in exploring new ideas and the generous funding to do so. Special thanks to Dr. Tonsor for continuing to provide guidance even after relocating to Kansas State University, and for challenging me to reach my full potential by never giving easy answers. Appreciation goes to Dr. Howard for always offering alternative viewpoints and for inspiring me to continue working in the area of sustainable food systems. Warm gratitude to Dr. Mariah Ehmke for providing me with valuable advice on life and career, and for setting an example that I aspire to.

I have made many wonderful connections within the food community at Michigan State University and in the Greater Lansing area during my graduate study. Specifically, I thank Julie Cotton for sharing her enthusiasm for local food systems and for the amazing work she is doing to expand opportunities for students interested in sustainable agriculture at MSU. I also thank all of the beautiful souls at MSU's Student Organic Farm, especially the 2010 Farmer Training Program cohort, for shared meals and showing me first hand what farm-to-table is really all about. A huge outpouring of love to Megan Sheahan, Rachel Ford, Miriam and Juha Sohlberg, and Ty Lawson for their delicious food, adventures, and friendship, without which I never would have made it through.

Finally, I would like to thank my family, both Sackett and Brian/Kleinhardt, for their unending support and love. A very special thanks to my wife, Rachel, for her constant companionship, for keeping me nourished in every sense of the word, for helping me keep things in perspective, and for working towards our shared dreams.

TABLE OF CONTENTS

LIST OF TABLES

KEY TO ABBREVIATIONS

GC = Ground Cover and Area Management Practices are Employed

FN = Fertilizer and Nutrient Materials are used minimally

CP = Little to No Chemical Pesticides are Used for Pest Management

CH = Little to No Chemical Herbicides are Used for Weed Management

PM = Pollinator Management is Employed

OP = Other Pests are controlled using preventative measures, and habitat controls

SA = Prohibited use of sub-therapeutic antibiotics

GH = Prohibited use of growth hormones

GM = Prohibited use of genetically modified livestock

AS = Animal Health and Safety are Protected

PF = Feed is Pasture Based and Waste Management Systems Employed

FS = Farm Size is small and Corporate Involvement is limited

PL = Production, distribution, and sale is done locally

CFP = Consumer food prices are affordable

FF = Farmers are financially stable

Chapter 1: General Introduction

Overview

Food produced using sustainable production practices is receiving increasing attention as the public and private sectors become progressively more invested in "sustainability" and "green marketing". There are currently over 200 different seals and logos representing ecological, ethical, or sustainability attributes in the global food industry (Organic Monitor). Alternative agricultural production systems generate value-added food attributes that provide consumers with varying degrees of utility, stemming from their environmental, economic, and social impacts. If farms are to remain viable and contribute to food and farm system sustainability, they may need to find and exploit these kinds of high-valued niche markets. Moreover, a firm's ability to differentiate their product depends critically on an accurate understanding of how such an alternative credence labeling scheme is perceived by the consumer. Building on existing work evaluating food attribute labels, this research examines the impact of consumer preferences on purchases of food products making "sustainably produced" claims.

As sustainable food claims gain market momentum, questions naturally arise regarding which attributes of sustainability are most important to consumers, how perceptions of sustainability may affect purchasing decisions, if consumer segments can be identified that systematically behave differently in response to these labeling schemes, and if purchasing decisions over specific sustainability attributes are stable and consistent in hypothetical and non-hypothetical settings. The limited existing literature on sustainable food labels has spent little effort identifying which attributes of sustainability consumers believe are important components of agricultural production systems. Additionally, because there is no government agency overseeing certification of these production-level claims, evidence suggests that consumers may be confused about the methods underlying the "sustainably produced" label.

Objectives

The primary objectives of this research are as follows: first to identify what consumers believe "sustainably produced" means by determining the aggregate importance and share of preference for a variety of sustainable farming practices considered. Second, we aim to estimate willingness to pay for sustainably labeled foods and compare estimates across hypothetical and non-hypothetical preference elicitation. Additionally, we seek to determine the impact of information on sustainability, from different sources, on WTP. Lastly, we develop a set of recommendations, based on the results of these objectives, designed to guide meat and produce industries in the development and implementation of a "sustainably produced" marketing plan.

To achieve these goals, we gathered a comprehensive set of information from consumers, through two complementary activities: a nation-wide survey on food consumption habits, including simulated shopping scenarios designed to reveal tradeoffs perceived by the consumer and a valuation field experiment designed to capture the same information in an in-person, non-hypothetical setting. The juxtaposition of the hypothetical and non-hypothetical results enhances the reliability and appropriateness of the economic analysis and marketing recommendations. Collectively, these two approaches yield a more detailed understanding of consumer behavior.

Literature Review

Surprisingly little economic research has focused on sustainability in the context of agricultural production practices. However, some food attributes that consumers may associate with sustainable production, such as local, organic, natural, GMO-free, hormone-free, and environmentally friendly have a considerable research base. Generally, the literature has found that credence attributes have a positive impact on consumer valuations and lead to higher WTP, despite the inability to directly identify the attribute either before purchase or after consumption. "Sustainably produced" is an attribute with no absolute definition, leav-

ing sustainable food labels much more open to consumer perceptions and inferences that go beyond the information provided on packaging or marketing labels. Understanding the perceptions consumers hold regarding what a "sustainably produced" labeling scheme implies is crucial to determining if the product will succeed in the market.

Preliminary Findings

Likert Scale Sustainability Perceptions

To collect consumer data used for this analysis, a national web-based survey of 1002 households was disseminated in the summer and fall of 2010. Consumer respondents were recruited by Decipher, a marketing research and survey programming company, and roughly adhere to the expected demographic profile of the United States population. The following Likert scale questions were posed to establish general perceptions of sustainability prior to completion of the market choice-oriented questions.

Table 1: Likert Scale Sustainability Perceptions

	Likert Scale* Ranking Share						
Survey Question	1	2	3	4	5	6	7
"I rarely think about the sustainability of production practices when making food purchases"	0.06	0.07	0.12	0.22	0.18	0.19	0.16
"Farmers face lower levels of productivity when using sustainable production practices"	0.04	0.09	0.15	0.39	0.18	0.11	0.05
"Farmers could sustainably produce food without much additional monetary expense"	0.05	0.09	0.16	0.33	0.20	0.10	0.07
"Farmers currently participate in sound sustainable food production practices"	0.02	0.05	0.11	0.39	0.24	0.13	0.07
"Organic products are from farms currently participating in sustainable practices"	0.03	0.04	0.09	0.31	0.21	0.21	0.10
"Locally grown products are from farms currently participating in sustainable practices"	0.02	0.04	0.08	0.38	0.24	0.16	0.09
"Smaller farms are more likely to currently participate in sustainable practices"	0.02	0.04	0.08	0.31	0.25	0.22	0.10
"Corporate farms are more likely to currently participate in sustainable practices"	0.07	0.11	0.17	0.33	0.16	0.11	0.06
"All products labeled as sustainably produced are from farms currently participating in sustainable practices"	0.03	0.04	0.13	0.38	0.21	0.15	0.07

*Scale is measured according to 1 = Strongly Disagree and 7 = Strongly Agree

Locally Grown Perceptions

When asked, " Which best describes the proximity from your home you consider locally grown products to come from?", the survey participants responded as follows:

Table 2: Locally Grown Perceptions

Distance	Choice Share
10 miles	0.17
50 miles	0.30
100 miles	0.20
500 miles	0.07
Within My State	0.24
Other	0.02

Social Issues Importance

Finally, when asked to rank the following social issues in order of their perceived importance, the participants responded:

Table 3: Social Issues Importance

Social Issue	Mean Ranking	Share of #1 Ranking
Food Safety	3.13	0.26
Human Poverty	3.88	0.20
Health Care	3.99	0.18
Consumer Food Prices	4.11	0.14
Environment	4.5	0.11
Animal Welfare	5.63	0.05
Sustainable Agriculture	5.45	0.03
Financial Stability of Farmers	5.3	0.02

Organization of Dissertation

This dissertation proceeds as follows, beginning with an examination of consumer perceptions of select sustainable farming practices in Chapter 2. The second chapter begins to address gaps in the literature regarding food products with sustainable production claims through the implementation of a best-worst scaling framework. The goal of the best-worst analysis

is to identify what consumers believe "sustainably produced" should mean in the context of food labels. The primary purpose of Chapter 2 is to introduce an economic application of best-worst scaling for measuring importance of production attributes in food systems, to identify which environmental, economic, and social indicators of sustainability currently used by third-party agricultural certifiers as farm-level evaluation criteria consumers perceive as important systematic components of food production, and to assess consumer heterogeneity for insight into potential marketing strengths for third-party certified "sustainable" products.

The dissertation continues to unpack preferences over sustainable food labels utilizing a hypothetical choice experiment in Chapter 3. The third chapter seeks to estimate preferences for "sustainably produced" food products in a discrete choice modeling framework and to determine how preferences may be affected by varying degrees of information about sustainable agricultural systems from different sources. Using comparative measures, Chapter 3 begins to shed light on the degree of perceived market differentiation between products labeled as "sustainable" and their "organic" and "local" counterparts.

Building further on the analysis from Chapter 3, the fourth chapter in this volume examines the hypothetical and non-hypothetical stated preferences of a smaller sub-sample of consumers that participated in experimental auctions in a laboratory setting. Comparison of the hypothetical and non-hypothetical WTP measures is used to determine the extent of hypothetical bias in the sample, and to induce a relationship between hypothetical choices and bidding behavior. Together these two measures, analyzed in Chapter 4, provide a richer description of consumer preferences and is believed to be more representative of actual market behavior. In Chapter 5 general conclusions are drawn and a synthesis of the results leads to a set of recommendations for producers and marketers of sustainably produced foods.

Chapter 2: Consumer Perceptions of Sustainable Farming Practices

Introduction

Food produced using sustainable production practices is receiving increasing attention in both public and private arenas as a greater number of food products are being marketed and labeled using "sustainable" or "sustainably produced" certification schemes for differentiation. As sustainably produced food gains market momentum, questions arise regarding what consumers perceive when faced with "sustainably produced" labels. Specifically, consumers want to know what sustainable claims imply about the environmental, economic, and social factors of production and farmers want to know what consumers are willing to pay for this value-added attribute before either party invests heavily in sustainable certification.

Consumer desired attributes of food system sustainability have recently been discussed in more detail in the literature. For example, a framework for evaluating consumer priorities with regard to sustainable foods was built by Clonan et al (2010) based on seven guiding principles of sustainability put forth by Sustain, an alliance for better food and farming. The authors used a five point Likert scale, embedded in a structured questionnaire, to explore attitudes towards sustainability components such as fair trade, organic, local, and animal welfare. This study found that consumers responded positively toward environmental responsibility metrics related to how their food was produced. Similarly, using a discrete choice modeling method for evaluating consumer attitudes towards sustainability claims on food products, Saunders et al (2010) also focus on the results of a Likert scale rating of sustainability attributes, in the context of carbon emissions and other contributions to global climate change. However, the limited literature on sustainable food labels has spent very little effort focused on identifying attributes of sustainability that consumers believe are or should be important components of a "sustainably produced" labeling scheme.

Batte(2010) reviews several studies that identify consumer-driven changes in food marketing channels as related to sustainable food claims. In his review, Batte identifies three

studies that support the importance of consumer demand in food product differentiation schemes. The first, a study by Onozaka et. al. (2010) finds evidence from a conjoint choice experiment of significant heterogeneity in valuing various food differentiation claims among shoppers in different marketing venues and notes that the consistent significance of self-perceived efficacy in their psychographic model suggests that consumers that believe they have a role in improving sustainability tend to value sustainable product claims more highly. In another study, Onozaka and McFadden (2011) explore the increasing use of sustainable food labels by analyzing the interactive effects of sustainable production claims, finding that locally grown differentiation is most highly valued. The authors suggest that consumer preferences for local food have been found to go beyond basic quality characteristics and are increasingly significantly related to sustaining the local economy by supporting local farmers and conserving local farmland. Batte concludes that further research is needed to identify how consumer demand for sustainably produced food is affected by the perceived importance of the environmental, economic, and social attributes used in differentiation through certification and labeling.

To address this gap in the literature, we utilize a best-worst scaling framework to examine what consumers believe "sustainably produced" - in the context of food labels - means or should mean. Best-worst choice modeling is a relatively new technique for analyzing consumer preferences or beliefs. This technique has been applied in a variety of settings including public health research by Flynn et all (2007), international agribusiness marketing by Umberger et al (2010), and by Scarpa et al (2010) to estimate desired tourism benefits. In an agricultural and food systems context, best-worst choice design and analysis have been previously demonstrated using empirical wine market examples by Casini, Corsi and Goodman (2009), Cohen (2009), and Mueller and Rungie (2009). More closely related to our work, Lusk and Parker (2009) use best-worst scaling techniques to evaluate the methods consumers prefer for producers to use to improve the fat content in ground beef. Most closely related to the current study, Lusk and Briggeman (2009) determine relative consumer

attitudes towards value-added food attributes, such as safety, nutrition, taste, and price using the best-worst framework. Here we examine consumer attitudes towards value-added sustainable production attributes in a similar fashion.

The primary purpose of this paper is threefold. First, to introduce an economic application of best-worst scaling for measuring the importance of production attributes in consumer decision making in the context of food systems. Additionally, we seek to identify which environmental, economic, and social indicators of sustainability, currently utilized by third party agricultural certifiers as farm/ranch-level evaluation criteria, consumers perceive as important systematic components of food production. Finally, we aim to assess consumer heterogeneity in perceived importance of sustainable farming practices when facing "sustainably produced" food purchasing decisions for insight into potential marketing strengths for third party certifiers.

Research Methodology

Best-Worst Scaling

Marketing surveys that measure attribute importance most often utilize a Likert Scale ranking approach. However, this method has several known weaknesses. First, scaled rating systems do not force respondents to make trade-offs between attributes. Additionally, Likert Scale ranked data lack a natural interpretation outside of the survey context. To address these issues, we implemented a best-worst design to investigate preferences for and perceptions of alternative sustainable farming practices. The survey instrument used to collect consumer data was designed to simplify the choice task for respondents.

Best-Worst analysis requires the survey respondent to simultaneously choose the most and least important attributes out of a set of competing options. This method is commonly referred to as "maximum difference scaling", as the attributes chosen should maximize the difference in utility realized, for the respondent, on an underlying scale of preference. The measured level of importance from the best-worst data analysis is applied to a standardized

ratio scale that determines, with more certainty, the percentage difference in importance across attributes. The theoretical foundation for this analysis is provided by Marley and Louviere (2005) in their development of probabilistic models for analyzing best-worst choice tasks.

Best-worst scaling, as originally devised by Flynn and Louviere (1992), is capable of addressing relative utility impacts across attributes that traditional discrete choice questions cannot. To observe trade-off behavior, the specification of attributes from a choice set of competing alternatives is repeated over a number of variable choice subsets. In this way, best-worst tasks provide more information than single choice designs, while forcing respondents to consider the extremes of their utility space. Ideally, the stated preference outlined in each best-worst scenario should approximate observed consumer behavior in retail markets.

Each choice set was created using a 2^{10} main effects orthogonal experimental design that was balanced with ten attributes, each exhibiting two levels. The specific attributes chosen for inclusion are motivated in the following subsection and further outlined in the Data Collection section. The orthogonal experimental design yielded twelve alternative choice sets, broken into two blocks and randomized across participants. Each block contained two questions with five alternatives, three questions with six alternatives, and one question with all ten alternatives. The presence or absence of each sustainable farming attribute was independent across choice subsets, allowing identification of relative preferences onto a ratio scale. The additional utility or dis-utility from moving between attribute levels will be estimated using a logit model framework.

Sustainability Attributes

The USDA National Institute of Food and Agriculture provides limited information on the purported sustainability of particular farming and ranching practices. At the national level, sustainable agriculture was first addressed by Congress in the 1990 Farm Bill (USDA NIFA) which states that under the law,

"The term *sustainable agriculture* means an integrated system of plant and animal production practices having a site-specific application that will, over the long term:

- satisfy human food and fiber needs
- enhance environmental quality and the natural resource base upon which the agricultural economy depends
- make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls
- sustain the economic viability of farm operations
- enhance the quality of life for farmers and society as a whole."

However, the USDA also warns that, "guidelines about what specific practices meet long-term environmental, economic, and social goals and constitute sustainable agriculture is still under debate". Therefore, when developing our consumer survey, for the purposes of determining perceptions of the practical farm-level components of a sustainable production system, we attempted to best follow the advice provided by the USDA by including environmental, economic, and social attributes of sustainability that address the final two points above.

The USDA provides website url links to a handful of groups that have attempted to develop standards and provide certification services based on their own guidelines. The first of the groups indicated on the USDA website is Food Alliance ($www.foodalliance.org$). Food Alliance is a non-profit organization that has developed sustainable agricultural practice

standards and operates a voluntary certification program based on those standards. Food Alliance is the most comprehensive certification program for sustainable food production in North America, employing independent third-party inspectors to audit Food Alliance certified businesses to determine whether they meet program standards and criteria. Food Alliance outlines all crop and livestock specific certification criteria and provides the following general guidelines for all certified production:

- Provide safe and fair working conditions
- Ensure the health and humane treatment of animals
- No use of hormones or sub-therapeutic antibiotics
- No genetically modified crops or livestock
- Reduce pesticide use and toxicity
- Protect soil and water quality
- Protect and enhance wildlife habitat
- Continuously improve management practices

For the purposes of this study, we adhered as closely as possible to the Food Alliance Whole Farm/Ranch Evaluation Criteria for crop and livestock specific guidelines when choosing the sustainability attributes that would be included in the survey. All crop and livestock specific Food Alliance Evaluation Criteria is publicly available online and from the authors by request. Because there is currently no government sponsored certification of sustainable food production, all "sustainably produced" labels found on food products in the United States are certified by a private third-party such as Food Alliance, or by the farm of origin itself. We are interested in identifying which certification-guided environmental, economic, and social attributes consumers perceive as important indicators of sustainability when faced with these kinds of purchasing decisions.

Other reputable, broadly defined, sustainable agriculture certifying organizations include Rainforest Alliance and Global G.A.P. The Rainforest Alliance certification standards are based on the same Sustainable Agriculture Standards (SAS) as Food Alliance. However, the Food Alliance certification program was chosen as the guiding standard because it had more detailed crop and livestock specific criteria publicly available at the time the survey was developed. Food Alliance standards took preference over Global G.A.P guidelines because of its national versus international focus and because Global G.A.P functions to oversee other accredited certification bodies for a number of different labels, potentially creating confusion.

There are many farming practices with sustainable characteristics that could have been included for the purposes of this study. The set of sustainable farming attributes were chosen for inclusion after study of the USDA's general sustainability guidelines and Food Alliance's third party sustainable certification standards. The set of attributes used in this study offer insight into possible perceived dimensions of sustainability and could be expanded on in future work. It is worth note that due to the nature of the variable choice sets, the analysis that follows is conditional on the set of evaluated attributes/farming practices and should not be interpreted to make judgements on sustainability attributes outside of this context.

Data Collection

This is the first known study to use best-worst scaling to measure consumer perceptions of production attribute importance in the context of food system sustainability. We chose best-worst scaling because we are especially interested in determining relative importance of sustainable agricultural certification criteria currently used in food markets. To collect consumer data used for this analysis, a national web-based survey of 1002 households was disseminated in the summer and fall of 2010. Consumer respondents were recruited by Decipher, a marketing research and survey programming company. A summary of population demographic statistics can be found in the following table.

Table 4: Demographic Variables and Summary Statistics of Survey Participants

Variable	Definition	Apple	Beef
Gender	1= Male; 2 = Female	1.476	1.516
	Total Participants	500	502
Age	Average Age in Years	51.48	50.92
Adults	Number of Adults in Household	2.062	2.048
Children	Number of Children in Household	0.48	0.51
Meals	Number of Meals/ week with Product	6.64	3.19
Shop	% of Total Shopping at Location		
	Grocery Store	81.97	81.16
	Health Food Store	8.49	9.20
	Food Co-op	2.05	1.90
	Convenience Store	3.0	4.42
	Farmers Market	3.72	5.24
	Butcher	4.38	2.92
Assistance	1 = On Food Assistance; 2 = Otherwise		
	1 = On Food Assistance	13.2 %	11.95 %
	2 = Not on Food Assistance	86.8 %	88.05 %
Education	Highest Level Completed		
	1 = Did not graduate from high school	2.2 %	2.59 %
	2 = Graduated from high school, no college	17.8 %	18.12 %
	3 = Attended college, no degree earned	28.8 %	33.67 %
	4 = Attended college, associates or trade degree earned	15.4%	12.55 %
	5 = Attended college, Bachelors degree earned	24.6 %	22.52 %
	6 = Graduate or advanced degree earned	11.2 %	10.56 %
Household Income	Range of Pre-tax Income		
	1 = Less than $20,000	20.8%	19.12 %
	2 = $20,000 - $39,999	26.4%	28.88 %
	3 = $40,000 - $59,999	18.4 %	21.51 %
	4 = $60,000 - $79,999	17.6 %	11.95 %
	5 = $80,000 - $99,999	8.2 %	7.57 %
	6 = $100,000 - $119,000	3.8 %	4.98 %
	7 = $120,000 - $139,999	1.8%	2.39 %
	8 = $140,000 - $159,999	1.4 %	1.20 %
	9 = More than $160,000	1.6%	2.39 %

Two versions of the survey were developed; one for apples and one for beef. The following table lists the sustainable farming practices that were chosen, in accordance with the crop (apple) and livestock (beef) specific Food Alliance Evaluation Criteria, as the best-worst attributes that the survey respondents would choose between for each food product.

Table 5: Apple and Beef Survey Attributes

Apple Survey Attributes	Beef Survey Attributes
Ground Cover Management	Prohibited Use of Antibiotics
Limited Fertilizer Use	Prohibited Use of Growth Hormones
Limited Pesticide Use	Prohibited Use of GM Livestock
Limited Herbicide Use	Animal Health and Safety
Pollinator Management	Pastured Feed and Waste Management
Preventative Pest Control	Preventative Pest Control
Farm Size	Farm Size
Geographic Level of Production	Geographic Level of Production
Consumer Food Prices	Consumer Food Prices
Financial Stability of Farmers	Financial Stability of Farmers

Many of these practices fall under current organic certification guidelines endorsed by the USDA, and all aforementioned attributes are components of sustainable farm certification by Food Alliance. The included attributes span the three-pronged sustainability framework suggested by Callens and Tyteca (1999) and are supported by the USDA National Institute of Food and Agriculture, using environmental, economic, and social metrics for evaluation.

The choice sets for this analysis are developed around ten attributes, each with two levels indicating the presence or absence of a given farming practice. Consumers are shown a set of attributes and asked to indicate which is most important (best) and which is least important (worst) over several repeated choices where the set of attributes varies across questions. In theory, each respondent undertakes the task of identifying every possible pair of attributes, calculating the difference in utility between each attribute pair, and choosing the pair that maximizes this difference. This method is an extension of Thurstone's (1927) paired comparison method, which has been utilized frequently in psychological research. In maximum-difference scaling, the distances between attributes are modeled as pair-wise

utilities and estimated in relation to a single attribute level rather than to an entire scenario. All apple survey participants were faced with the following example scenario, with similar extension to the beef survey.

Survey Question Example

Which one of the following aspects of apple farming do you believe is the most and least important in a sustainable apple production system? Please check only one in each column.

Table 6: Best-Worst Survey Question Example

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Little to No Chemical Pesticides are Used for Pest Management	
	Pollinator Management is Employed	
	Other Pests are controlled using preventative measures, and habitat controls	
	Production, distribution, and sale is done locally	
	Consumer food prices are affordable	

The main effects orthogonal experimental design ensured that each potential best-worst pair appeared exactly twice in each blocked section of six choice sets and that each attribute was seen by the respondent exactly four times in this portion of the survey. Definitions were provided for six of the production attributes, excluding the economic attributes, which were left to the interpretation of each respondent. Full definitions can be found in the Appendix to this paper.

Analysis

Two primary approaches are used to analyze best-worst ranked data. "Paired" models are implemented to make inferences about the latent utility scale, while "marginal" models aggregate over all pairs that include a given attribute level, to model choice frequencies.

Both methods have the same measurement properties and can be analyzed at the respondent or sample level, yielding similar results. For brevity, the simpler marginal (count-based) specification and analysis is omitted in this text and is available in the Appendix to this chapter.

Consider a choice set like the one outlined above, with $J = 6$ attributes, and therefore $J(J-1) = 30$ possible best-worst combinations that could be chosen by the respondent. The particular pair of attributes chosen by the consumer represents one choice out of all $J(J-1) = 30$ possible pairs, in a given choice set, that maximizes the difference in their perceived importance. Let λ_j formally represent the location of attribute j on the underlying scale of importance. Therefore the latent, unobservable level of importance for individual i is given by

$$I_{ij} = \lambda_j + \epsilon_{ij} \tag{1}$$

where ϵ_{ij} is the random error. Thus, the probability that the consumer chooses item j as most important and item k as least important is equal to the probability that the difference between I_{ij} and I_{ik} is greater than all other $J(J-1)-1$ possible differences in the choice set. If ϵ_{ij} is distributed i.i.d then λ_j can be estimated for attribute j in a conditional logit framework. The equation to be estimated is as follows:

$$choice_i = \alpha + \lambda_1 L_1 + \lambda_2 L_2 + ... + \lambda_J L_J + \epsilon_i \tag{2}$$

where each L_j corresponds to one of the J attributes included in a given choice set. Each attribute is an alternative specific regressor, taking a value in $-1, 0, 1$.

Using the conditional logit parameter estimates, λ_j, the probability that attribute j is chosen as most important (best) and attribute k is chosen as least important (worst) is given by:

$$Prob(j = best \cap k = worst) = \frac{e^{\lambda_j - \lambda_k}}{\sum_{s=1}^{J} \sum_{t=1}^{J} e^{\lambda s - \lambda t} - J} \tag{3}$$

The parameter estimates represent the share of importance for the given attribute, relative to the attribute ranked least important, normalized to zero. To return results consistent with standardized ratio scaling techniques, the share of importance for each attribute, equal to the forecasted probability of being chosen as most important, takes the following form:

$$share_j = \frac{e^{\lambda_j}}{\sum_{i=1}^{J} e^{\lambda_i}} \tag{4}$$

The preference shares must sum to one across all ten sustainable farming attributes. The above equation should be interpreted as the importance of attribute j on a ratio scale. The share of preference for a given attribute reflects both the true importance of the attribute and the relative uncertainty in the importance consumers place on the attribute. However, this specification neglects consumer heterogeneity, as the estimation requires the assumption that all individuals in the sample must place the same level of importance on each value, once they are ranked.

To more deeply explore consumer heterogeneity, a latent class cluster analysis is also employed. This clustering technique assumes that individuals belong to one of a predetermined number of latent classes. The Bayesian Information Criterion (BIC) determines the optimal number of latent clusters. This method involves minimizing within-cluster variance and maximizing across-cluster variance. The probability of membership in a given latent class is estimated using the logit model parameters. The latent class cluster analysis uses the individual best-worst scores as the dependent variables in this model. The covariation across individual best-worst scores measures unobserved utility gains and losses associated with the inclusion or exclusion of the given attribute for each respondent.

Results and Discussion

The raw data contains 1002 observations (rows of data), one for each respondent. Within each respondent's observation, two variables for each question corresponds to the respondent's choice of most important production attribute and choice of least important production attribute, leading to 12 variables for each respondent because each survey contains six best-worst questions. This type of data is commonly referred to as being in "wide-form". Using SAS, the wide-form data was transformed into what is commonly referred to as "long-form" data, with one observation providing data for just one alternative for each individual.

Specifically, if a given best-worst question offers J attributes to choose between, then the long-form data has $J(J-1)$ observations for that question, as alternatives are now thought of as potential best-worst pairings among the J potential attributes. Each observation for a given question includes J explanatory variables, one for each of the production attributes included in a choice set. In a given observation, all but two of the production attributes take a value of zero. Of the two remaining attributes in this observation, one production attribute takes a value of 1 referring to the "most important" choice, and the other takes a value of -1 referring to the "least important" choice. Thus, the $J(J-1)$ observations each refer to a possible pair of production attributes that could be chosen by the respondent as best and worst simultaneously. Over these $J(J-1)$ observations, the variable "choice" takes a value of zero for all but one observation. The remaining observation refers to the actual best-worst pair chosen by the respondent for that question and the variable "choice" takes a value of 1.

Wide-Form Data Example

Table 7: Wide-Form Data Example

ID	Q15 Options	Q15 Best	Q15 Worst
1	5	1	2
2	5	3	2
3	5	5	4
4	5	3	1

In this example, there are four survey respondents referred to by their ID numbers ranging from $1 - 4$. "Q15 Options" refers to the number of attributes that are contained in choice set (question) 15 and is the same for all respondents. "Q15 Best" refers to which of the five attributes the respondent chose as "most important" in choice set (question) 15 and similarly, "Q15 Worst" refers to which of the five attributes the respondent chose as "least important" in choice set (question) 15. Here, respondent 3 has answered question 15 by choosing attribute 5 as "most important" and simultaneously choosing attribute 4 as "least important".

Long-Form Data Example

Table 8: Long-Form Data Example

ID	Choice	Q15-1	Q15-2	Q15-3	Q15-4	Q15-5
3	0	-1	1	0	0	0
3	0	-1	0	1	0	0
3	0	-1	0	0	1	0
3	0	-1	0	0	0	1
3	0	0	-1	1	0	0
3	0	0	-1	0	1	0
3	0	0	-1	0	0	1
3	0	1	-1	0	0	0
3	0	0	0	-1	1	0
3	0	0	0	-1	0	1
3	0	1	0	-1	0	0
3	0	0	1	-1	0	0
3	1	0	0	0	-1	1
3	0	1	0	0	-1	0
3	0	0	1	0	-1	0
3	0	0	0	1	-1	0
3	0	1	0	0	0	-1
3	0	0	1	0	0	-1
3	0	0	0	1	0	-1
3	0	0	0	0	1	-1

In the table above, a portion of the wide-form example has been transformed into long-form. As noted above, respondent 3 is presented with five sustainable farming production attributes in question 15. As such, there are twenty $(5 * 4)$ possible best-worst pairs that this respondent could choose, corresponding to the twenty observations (rows of data). In each row, each of the five production attributes, labeled Q15-1 through Q15-5 for simplicity, take a value in $\{-1, 0, 1\}$. For example, the first row in this table refers to the possibility that in Q15, attribute 1 is chosen as "least important" and attribute 2 is simultaneously chosen as "most important". The second row in this table refers to the possibility that in Q15, attribute 1 is chosen as least important and attribute 3 is simultaneously chosen as "most important", and so on. We see from the wide-form data above that respondent 3 answered question 15

by choosing attribute 5 as "most important" and attribute 4 as "least important". In the long-form data this choice is reflected in the variable "choice" taking a value of 1 for the observation corresponding to that specific best-worst pair and zero everywhere else.

Conditional Logit

The long-form data is used in a logit-model analysis here. The summary statistics from a conditional logit model analysis can be found in the following table for apple production and beef production, respectively. The regression parameters are transformed into preference shares, which offer more intuitive interpretation. The preferences shares sum to one, and each represents the proportional share of importance for the given attribute relative to consumer food prices, which is normalized to zero.

Half of the respondent sample was randomly assigned to complete the survey on apple production practices. The results in Table 3 indicate that the attribute corresponding to "farm size is small and corporate involvement is limited" has the highest preference share. As noted earlier, the preference share parameters report the importance of a given sustainable farming production attribute on a ratio scale, reflecting both the true importance of the attribute as well as relative uncertainty conveyed as the probability that an attribute is picked as more important than any other. Following farm size, pollinator management received the second highest preference share. Worth noting is the moderately high preference share on the attribute corresponding to "production, distribution, and sale is done locally", with just under half the importance relative to farm size. "Consumer food prices are affordable" was chosen as least important more often in every scenario than any of the remaining attributes and was therefore set as the base category.

These results should also be interpreted in relative magnitude. For example, pollinator management is inferred to be roughly half as important in a sustainable agricultural system as farm size to this sample of consumers. In fact, in a sustainable agricultural system, farm size is found to be just more than two times more important than local production, distribution

and sale. The use of off-farm chemical inputs such as fertilizers, herbicides, and pesticides all ranked between three and four times less important to a sustainable agricultural system than farm size and only marginally more important than consumer food prices. Overall, this sample of consumers indicated that the four economic attributes of sustainability included in this study were located near the boundaries of their utility spectrum.

Table 9: Relative Importance of Sustainable Attributes: Logit Estimates

Apple Attribute	λ	Preference Share	Beef Attribute	λ	Preference Share
Farm Size	1.77* (0.05)	0.28	Farm Size	1.41* (0.05)	0.28
Pollinator Management	1.25* (0.06)	0.15	Local	1.09* (0.06)	0.17
Ground Cover	1.22* (0.06)	0.13	Preventative Pest Management	0.80* (0.06)	0.12
Local	1.08* (0.06)	0.12	Pasture Based Feed	0.47* (0.06)	0.09
Preventative Pest Management	0.90* (0.05)	0.09	Prohibited Antibiotics	0.43* (0.06)	0.08
Minimal Fertilizer	0.78* (0.06)	0.08	Farmer Financial Stability	0.34* (0.06)	0.08
Limited Herbicides	0.56* (0.05)	0.06	Prohibited GMO	0.23* (0.06)	0.07
Farmer Financial Stability	0.29* (0.05)	0.04	Prohibited Growth Hormones	-0.08* (0.05)	0.05
Limited Pesticides	0.14* (0.05)	0.03	Animal Health and Safety	-0.10* (0.05)	0.04
Affordable Food Prices	DROPPED	0.02	Affordable Food Prices	DROPPED	0.02

(*) and (**) significant at 0.01 and 0.05 levels, respectively
standard errors are reported in parentheses

The remaining half of the respondents were randomly assigned to complete the survey on beef production practices. The data in Table 5 reveals that, as in the apple survey, the attribute corresponding to "farm size is small and corporate involvement is limited" has the highest preference share. The attribute with the second highest preference share is "production, distribution, and sale is done locally", suggesting that the locality of meat production is more important in a sustainable agricultural system to consumers than the locality of apple production relative to the other included attributes. This difference could be attributable to the fact that beef production is less place-specific than apple production due to the more specific climate needs for apple growth. The ratio scaled values indicate that locality is between one half and three quarters as important as farm size. The attribute corresponding to "animal health and safety is protected" had the lowest preference share other than affordable consumer food prices. The use of growth hormones also had a relatively low preference share.

When standardized to the ratio scale, farm size is found to be seven times more important in a sustainable beef production system to consumers than animal health and safety. Yielding the same results as the apple version of the survey, the respondents of the beef survey ranked financial well-being of farmers and affordable consumer food prices in the bottom third on the underlying scale of importance.

Latent Class Model

To better qualify distinct consumer segments, this data was also analyzed under a latent class clustering framework. The ten best-worst attributes formed the dependent variables over the underlying probability distribution of latent class inclusion. Consumer clusters with similar perceptions of the ten sustainable attributes are located close to one another in n-dimensional utility space. Identifying groups of consumers with similar perceptions is useful for advertising, pricing, and product development.

The optimal number of latent classes is determined by minimizing the Bayesian Informa-

tion Criterion. This resulted in four latent classes of apple consumers and five latent classes of beef consumers. The following tables outline the latent class structures. In each column, the parameter reported identifies the item response probability corresponding to participant endorsement of each attribute as "most important".

Table 10: LCA Estimates: Apple Attribute as "Most Important" by Cluster

Attribute	"Localvores" (14%)	"Small Farmers" (30%)	"Price Savvy Shoppers" (10%)	"Confused/ Indifferent" (46%)
Ground Cover	0.0	0.1852	0.0	0.0976
Fertilizer	0.0	0.1852	0.0	0.0976
Pesticides	0.0	0.1852	0.0	0.0976
Herbicides	0.3077	0.0	0.0	0.1219
Pollinators	0.0	0.1481	0.0	0.0976
Pest Management	0.3846	0.0	0.0	0.1220
Farm Size	0.0	0.1481	0.0	0.1219
Local	0.3077	0.0	0.0	0.1220
Food Prices	0.0	0.0	1.00	0.0
Farmer Stability	0.0	0.1481	0.0	0.1220

For apples, the first cluster, representing 14% of the sample, is made up of consumers that believe local production is a highly important aspect of sustainable production. This group also strongly values preventative pest management and limited use of chemical herbicides. It is possible that the use of chemicals in production is undesirable for this cluster due to potential effects on local ecosystems and water resources.

The second cluster, representing 30% of the sample, perceives limited corporate involvement to be important in a sustainable production system, as indicated by their preference for small farm size and financial stability of farmers relative to other clusters. This group also emphasizes the importance of other farming attributes that are in line with current land-use organic standards, such as the use of ground cover and the limited use of chemical inputs. This group likely places a higher value on organically labeled products, but would like to see the farms less consolidated in a sustainable system.

The third cluster, representing 10% of the sample, is motivated primarily by consumer

food prices. The fourth cluster, representing the largest portion of the sample at 46%, is made up of consumers that are either indicating indifference across these particular attributes, or may be confused about how these attributes relate to one another in a systematic way. This is evidenced by the roughly equal ranking of all attributes in importance to a sustainable production system. From this latent class assessment we find that 56% of the evaluated apple consumer sample is either price driven or largely indifferent to varied aspects of sustainability.

Table 11: LCA Estimates: Beef Attribute as "Most Important" by Cluster

Attribute	"Animal Rights Activists" (15%)	"Nutrition Buffs" (24%)	"Price Savvy Shoppers" (10%)	"Confused/ Indifferent" (41%)	"Say No to GMOs" (10%)
Antibiotics	0.0	0.1818	0.0	0.1351	0.0
Growth Hormone	0.0	0.2273	0.0	0.1081	0.0
GMO	0.0	0.0	0.0	0.0	1.0
Animal Safety	0.3077	0.0	0.0	0.1351	0.0
Pest Mngmt	0.0	0.0	0.0	0.1081	0.0
Pastured Feed	0.3846	0.2273	0.0	0.1081	0.0
Local	0.3077	0.0	0.0	0.1351	0.0
Food Prices	0.0	0.0	1.0	0.0	0.0
Farm Size	0.0	0.1818	0.0	0.1351	0.0
Farmer Stability	0.0	0.1818	0.0	0.1351	0.0

For beef, the first cluster, representing 15% of the sample, values the ethical treatment of animals in meat production, evidenced by the high propensity for choosing the preservation of animal health and safety, and pasture based feed as most important in a sustainable system. Additionally, this cluster perceives local production and limited corporate involvement as desired contributions to the sustainable system.

The second cluster, comprising 24% of the sample, is primarily concerned with the safety and nutritional aspects of their food decisions. This cluster most often chose the prohibition of sub-therapeutic antibiotics and bovine growth hormone along with pastured feed as the

most important attribute of sustainable production systems. Additionally, this cluster of consumers values small farm size and the financial stability of farmers. Since small farms may be inclined to have a more direct connection with their consumer base, better food safety and traceability standards might be perceived as attainable.

The third and fourth clusters share the same characteristics as the "price savvy shoppers" and "sustainably indifferent" consumers from the apple survey, at 10% and 41%, respectively, together comprising 51% of the beef sample. However, in the beef survey a distinct fifth cluster emerges, comprised of consumers that chose the prohibition of genetically modified livestock as most important in every choice set in which it appeared. While this segment of the market is small, it may offer significant opportunities for producers that do not use genetically modified products and for marketing firms that could emphasize this attribute in a labeling scheme.

Conclusions

Best-worst analysis was applied in this research to investigate the degree of importance consumers give to ten sustainable farming production attributes and in particular was used to determine behavioral differences across clustered subgroups of the population sample. The advantages of this methodology compared to more traditional stated preference analysis is evident in its higher discriminatory power for measuring trade-off decisions and in its wider applicability and interpretation outside of the survey context. While avoiding common rating bias, best-worst analysis results can be used in cross-national and cross-regional comparison studies on diverse populations and their judgment of similar attributes. This study gives credence to the strength of the best-worst method in yielding clear and simple interpretations. The simplicity of this analysis can be applied by marketing managers to gain insight into the evaluation behavior of different consumer segments for targeted food labeling.

The information gathered here from consumer data on perceptions of sustainable farming practices holds large potential for marketing managers. The unique best-worst framework

provides greater insight into determinants of market behavior than more commonly used Likert scale ranking approaches. In both beef and apple surveys, consumers indicated a strong perceptive correlation between sustainability and the size and locality of the farm of origin. This analysis suggests, similar to the findings of Onozaka and McFadden (2011), that consumers perceive quality difference for locally grown and distributed products. Supporting studies, such as that of Bond, Thilmany and Keeling-Bond (2008) give evidence that preferences for local food products are significantly related to factors affecting farmer viability, sustaining local farm land, and contributing to smaller, local economies. Our work supports these findings that scale and geographic range factor heavily into consumer perceptions of sustainably labeled food products.

Increasing attention drawn towards smaller, local farms has been a response to the widening awareness of global food system business conduct. Distrust has been growing for imported foods, especially meat products, with recent publicity on country of origin labeling requirements and other high profile food contamination cases. Sustainability claims on food targets many dimensions of consumer utility from quality and safety concerns to more intrinsic valuation connected to underlying food values such as fairness and environmental impact. In effect, some sustainability claims may be seen by consumers as substitutable, while others complementary; another point emphasized by Onozaka and McFadden (2011). The value of sustainable certification may only contribute marginally to the localness of a food product, while in other situations it may enhance commitment to a more well-rounded sustainable farming view point.

While consumers are generally familiar with organic standards outlining land-use and environmental impact variables of production, this study indicates that environmental indicators of sustainability are less important to consumers than economic dimensions. Based on our initial results, size, scale, and geographic scope capture the attributes of sustainability that are most important to consumers. Therefore, differentiating food claims on the level of locality provides a marketing avenue worth exploring further, as supported by the growing

literature on the local foods movement.

Our study also provides supporting evidence that the term "sustainable" could be causing more confusion than it is adding value in credence labeling schemes. The latent class assessments on both beef and apple consumer samples indicated that 50 - 55% of the evaluated population is highly price-driven or significantly indifferent to varied sustainability attributes. These results could be interpreted to suggest confusion or indifference, but they could also imply that consumers view sustainability as a bundle of attributes and that each alone is not sufficiently important to determine the sustainable nature of an agricultural system. Therefore, this study may support claims that only an integrated system of some or all of these attributes can be considered sustainable, as perceived by consumers. Given the associated heterogeneous economic welfare impacts that would come as a result of across-the-board market adjustments, such as government endorsed sustainability labeling, or bans on select farming practices, this is an important result.

APPENDICES

Marginal Analysis

Two primary approaches are used to analyze best-worst ranked data. "Paired" models are implemented to make inferences about the latent utility scale, while "marginal" models aggregate over all pairs that include a given attribute level, to model choice frequencies. Both methods have the same measurement properties and can be analyzed at the respondent or sample level, yielding similar results. For brevity, the simpler marginal (count-based) analysis specification is omitted in the published text.

For marginal (count-based) data analysis, individual best-worst scores are calculated for each attribute as the summation of the number of times each respondent indicated the attribute as most important less the summation of the number of times each respondent indicated the attribute as least important. The larger the best-worst score, the higher the attribute is ranked on the underlying scale of importance. Individual best-worst scores, for each attribute, are aggregated across the sample. The square root of the aggregate frequency of "chosen as best" divided by the aggregate frequency of "chosen as worst" is used to standardize scores to a ratio scale. This simple calculation is merely an empirical model-free estimate of the random utility model parameters. The highest $\sqrt{\frac{BEST}{WORST}}$ is scaled to 100 accordingly, and all other attributes scaled relative to this attribute.

The ratio scaled value for each attribute can be interpreted as the probability that the average respondent believes the given attribute is most important relative to the remaining $J-1$ attributes. The count-based, standardized ratio scale ranks the attributes in the same order as all other preliminary measures of attribute importance. In fact, the natural log of this quantity is a good estimate of the conditional logit model parameters ran on the same data. We should note here that the ratio scaled probabilities do not add to one because each choice scenario only contains a subset of the total choice set of ten attributes. Therefore, the probability of being chosen as most important is dependent on the inclusion of the other attributes in the choice scenario.

Table 12: Apple Attribute Importance Measures using Best-Worst Scaling

Attribute	Most	Least	Agg B-W	SQRT B/W	Std Ratio	Mean B-W	SD B-W	Ind Ratio
Farm Size is small and Corporate Involvement is limited	759	120	639	2.51	100	0.64	1.61	100
Pollinator Management is Employed	388	136	252	1.69	67.16	0.23	0.87	35.94
Ground Cover and Area Management Practices are Employed	370	138	232	1.64	65.11	0.23	0.91	35.94
Production, distribution, and sale is done locally	392	229	163	1.31	52.02	0.16	1.07	25
Other Pests are controlled using preventative measures, and habitat controls	247	193	54	1.13	44.98	0.05	0.77	7.81
Fertilizer and Nutrient Materials are used minimally	196	205	-9	0.98	38.88	-0.01	0.79	-1.56
Little to No Chemical Herbicides are Used for Weed Management	223	341	-118	0.81	32.15	-0.12	0.93	-18.75
Farmers are financially stable	135	438	-303	0.56	22.07	-0.3	1.09	-46.88
Little to No Chemical Pesticides are Used for Pest Management	162	531	-369	0.55	21.96	-0.37	1.13	-57.81
Consumer food prices are affordable	140	681	-541	0.45	18.03	-0.54	1.32	-84.38

Half of the respondent sample was randomly assigned to complete the survey on apple production practices. The results indicate that the attribute corresponding to "farm size is small and corporate involvement is limited" has the highest best-worst score, as well as being chosen as "most important" with the highest frequency. The ratio scaled parameter on the farm size attribute infers that any given respondent would choose it as most important with a probability of 0.64. Following farm size, pollinator management was ranked with a probability of choice of 0.23. Worth noting is the relatively high best-worst score on the

attribute corresponding to "production, distribution, and sale is done locally", with about one fourth of the importance relative to farm size. "Consumer food prices are affordable" received the lowest best-worst score and was chosen as least important more often in every scenario than any of the remaining attributes.

These results should also be interpreted in relative magnitude. For example, pollinator management is inferred to be roughly one third as important in a sustainable agricultural system as farm size to this sample of consumers. In fact, in a sustainable agricultural system, farm size is found to be four times more important than local production, distribution and sale, and is five times more important than affordable consumer food prices. The use of off-farm chemical inputs such as fertilizers, herbicides, and pesticides all ranked between three and four times less important to a sustainable agricultural system than farm size and only marginally more important than consumer food prices. Overall, this sample of consumers indicated that the four economic attributes of sustainability included in this study were located on the boundaries of their utility spectrum.

Table 13: Beef Attribute Importance Measures using Best-Worst Scaling

Attribute	Most	Least	Agg B-W	SQRT B/W	Std Ratio	Mean B-W	SD B-W	Ind Ratio
Farm Size is small and Corporate Involvement is limited	729	117	612	2.5	100	0.61	1.53	100
Production, distribution, and sale is done locally	519	119	400	2.09	83.66	0.4	1.06	65.57
Pests are controlled using preventative measures, cultural and nutritional controls	351	153	198	1.51	60.68	0.2	0.95	32.79
Feed is Pasture Based and Waste Management Systems Employed	221	208	13	1.03	41.29	-0.01	0.74	-1.64
Prohibited use of sub-therapeutic antibiotics	216	227	-11	0.98	39.08	-0.01	0.84	-1.64
Farmers are financially stable	223	304	-81	0.86	34.31	-0.08	1.05	-13.11
Prohibited use of genetically modified livestock	238	374	-136	0.8	31.96	-0.14	1.08	-22.95
Consumer food prices are affordable	194	517	-323	0.61	24.54	-0.32	1.27	-52.46
Prohibited use of growth hormones	181	506	-325	0.6	23.96	-0.32	1.14	-52.46
Animal Health and Safety are Protected	128	475	-347	0.52	20.8	-0.35	1.09	-57.38

The remaining half of the respondents were randomly assigned to complete the survey on beef production practices. The data reveals that the attribute corresponding to "farm size is small and corporate involvement is limited" has the highest best-worst score and the highest frequency of choice as "most important" across all respondents. The attribute with the second highest best-worst score is "production, distribution, and sale is done locally", suggesting that the locality of meat production is more important in a sustainable agricultural system to consumers than the locality of apple production relative to the other included attributes. The ratio scaled values indicate that locality is roughly three quarters as important as farm size. The attribute corresponding to "animal health and safety is protected" had the lowest

best-worst score and was chosen as least important more times in every scenario than any of the remaining nine attributes. Affordable consumer food prices and prohibited use of growth hormones also had relatively low best-worst scores.

When standardized to the ratio scale, farm size is found to be more than five times more important in a sustainable agricultural system to consumers than animal health and safety. Yielding the same results as the apple version of the survey, the respondents of the beef survey ranked financial well-being of farmers and affordable consumer food prices in the bottom third on the underlying scale of importance. The use of sub-therapeutic antibiotics and pasture based feed and waste management systems both exhibit a near zero mean, indicating indifference among consumers.

The aggregated best-worst scores do not convey information about the heterogeneity that may be present across sample consumers. The larger the range of heterogeneity across consumers, the more the market will respond to targeted channels of communication through labeling schemes. The standard deviation of the best-worst score for each attribute measures the extent of variation amongst consumers. A higher standard deviation indicates a wider variety of relative importance for a given attribute. Conversely, a smaller standard deviation is indicative of general agreements across consumers on relative importance of a given attribute.

The standard deviation is bounded by [-4,4] because each attribute could be chosen as most or least important a maximum of four times in each version of the survey. For the purposes of this study, a standard deviation above one can be interpreted as high heterogeneity across consumers. Attributes at both ends of the spectrum of relative importance exhibit varying degrees of heterogeneity. Attributes with a high best-worst score and high standard deviation are likely to be very important to a select subset of consumers. Additionally, attributes with a low best-worst score and high standard deviation have potential in niche markets, since the attribute appeals to a much smaller consumer segment.

Distinct drivers of heterogeneity are important to identify which attributes are related or

jointly important for the same consumer segment. The variance-covariance matrix outlines the pairs of attributes that vary simultaneously. For example, attributes that are highly covariable will both exhibit a high best-worst score within the same consumer group. Similarly, attributes that are negatively correlated will likewise drive the same segment of consumers but to opposite conclusions on importance. For this reason, correlation coefficients are often easier for interpretation, due to their natural bounding in [-1,1]. The basis for the cluster analysis used in the body of this paper comes from attributes that tend to be tracked together over consumers. Therefore, greater numbers of statistically significant correlation coefficients imply a more structured market.

The highly negative correlation between farm size and consumer food prices indicates that these attributes are likely to move in opposite directions of importance within a given group of consumers. Another interesting result is the highly negative correlation between farm size and chemical pesticide use. Additionally, the attribute corresponding to local production, distribution and sale is significantly negatively correlated with both chemical pesticide and herbicide use. Together, these observations suggest that there may be two distinct consumer segments in this market; one that views traditional organic certification standards, outlining chemical use and land management, as important components of sustainability, and one that believes supporting local economic enterprise and small business is more important to the sustainable model of production. We should note that these two beliefs do not need to be in conflict, but they do offer insight into the perception of sustainability as it relates to more well known organic food standards and how the systems may or may not compliment one another.

Table 14: Apple Production Attribute Variance-Covariance Matrix

	GC	FN	CP	CH	PM	OP	FS	PL	CFP	FF
GC	0.83									
FN	-0.05	0.63								
CP	-0.25	0.04	1.28							
CH	-0.12	0.07	0.46	0.86						
PM	0.12	-0.01	-0.22	-0.12	0.75					
OP	0.09	-0.03	0.08	-0.02	0.02	0.59				
FS	-0.16	-0.26	-0.54	-0.32	-0.18	-0.27	2.58			
PL	-0.03	-0.13	-0.34	-0.25	-0.08	-0.16	-0.03	1.15		
CFP	-0.23	-0.13	-0.22	-0.22	-0.33	-0.11	-0.51	-0.09	1.76	
FF	-0.15	-0.09	-0.24	-0.28	-0.07	-0.15	-0.3	-0.06	0.14	1.19

Table 15: Apple Production Attribute Correlation Coefficients

	GC	FN	CP	CH	PM	OP	FS	PL	CFP	FF
GC	1									
FN	-0.06	1								
CP	-0.24	0.04	1							
CH	-0.15	0.1	0.45	1						
PM	0.15	-0.01	-0.22	-0.15	1					
OP	0.12	-0.05	0.1	-0.03	0.03	1				
FS	-0.11	-0.2	-0.31	-0.22	-0.12	-0.22	1			
PL	-0.04	-0.15	-0.28	-0.24	-0.09	-0.2	-0.01	1		
CFP	-0.2	-0.13	-0.15	-0.18	-0.28	-0.11	-0.24	-0.06	1	
FF	-0.15	-0.12	-0.19	-0.28	-0.07	-0.18	-0.17	-0.05	0.1	1

Table 16: Beef Production Attribute Variance-Covariance Matrix

	SA	GH	GM	AS	PF	OP	FS	PL	CFP	FF
SA	0.71									
GH	0.18	1.31								
GM	-0.09	0.28	1.16							
AS	-0.07	-0.1	0.02	1.19						
PF	-0.01	-0.04	-0.08	0.04	0.54					
OP	-0.05	-0.14	-0.09	-0.06	0.04	0.9				
FS	-0.27	-0.65	-0.35	-0.37	-0.13	-0.22	2.33			
PL	-0.05	-0.26	-0.27	-0.24	-0.15	-0.07	0.05	1.12		
CFP	-0.15	-0.29	-0.36	-0.25	-0.14	-0.11	-0.31	-0.12	1.6	
FF	-0.19	-0.27	-0.22	-0.12	-0.14	-0.18	-0.08	-0.01	0.13	1.1

Table 17: Beef Production Attribute Correlation Coefficients

	SA	GH	GM	AS	PF	OP	FS	PL	CFP	FF
SA	1									
GH	0.19	1								
GM	-0.09	0.22	1							
AS	-0.09	-0.09	0.02	1						
PF	-0.02	-0.05	-0.1	0.06	1					
OP	-0.05	-0.13	-0.09	-0.05	0.06	1				
FS	-0.21	-0.37	-0.21	-0.22	-0.12	-0.15	1			
PL	-0.05	-0.21	-0.24	-0.2	-0.19	-0.07	0.03	1		
CFP	-0.14	-0.19	-0.26	-0.18	-0.15	-0.09	-0.16	-0.1	1	
FF	-0.21	-0.21	-0.19	-0.09	-0.18	-0.18	-0.05	-0.02	0.09	1

The beef survey results highlight the high heterogeneity across consumers in their importance ranking of the use of growth hormones, genetically modified livestock, and the protection of animal health and safety. It is interesting to note that these three attributes also exhibited highly negative correlation with farm size. These three attributes hinge critically on food safety and human health and nutrition dimensions of consumer utility and may be motivated by intrinsic food values deeply connected to eco-responsibility movements and the emergence of socially alternative food markets. These results support the findings of Umberger, McFadden and Smith (2009) examining consumer valuation of hormone and GM free claims, leading us to the same conclusion that social dimensions of food values have the potential for creating distinct segments in the consumer market.

Apple Survey Attribute Definitions

In the next section you will be asked to choose which aspects of sustainable apple farming are most and least important to you. Please take the time to read the following definitions as related to sustainable production practices to better help you in your responses.

Ground Cover and Area Management Practices are Employed: Adjacent areas are planted with hedgerows, windbreaks, or other low-maintenance plantings to encourage specific beneficial organisms. Within tree rows, ground cover or mulch are selected and maintained to improve soil microbial activity, organic matter levels and nutrient cycling.

Fertilizer and Nutrient Materials are Used Minimally: Soil quality, including organic matter content, is established at planting and maintained at an optimum level to minimize commercial fertilizer needs.

Little to No Chemical Pesticides are Used for Pest Management: Chemical pesticides are not used. Alternative strategies are employed, including biopesticides, mating disruption, trap out and/or augmentation with beneficial organisms.

Little to No Chemical Herbicides are Used for Weed Management: Soil quality and ground cover in the orchard and adjoining areas are planned and managed to prevent weeds and weed seed immigration into the orchard. Cultural, mechanical or biological methods are used to control weeds.

Pollinator Management is Employed: Bees are not placed in the orchard until blossoms are open. Pesticides hazardous to bees are not used, or only if needed in an emergency.

Other Pests are Controlled Using Preventative Measures, and Habitat Controls: Habitat is modified around orchards to reduce nesting and perching sites for pest birds.

Beef Survey Attribute Definitions

In the next section you will be asked to choose which aspects of sustainable cattle farming are most and least important to you. Please take the time to read the following definitions as related to sustainable production practices to better help you in your responses.

Prohibited Use of Sub-therapeutic Antibiotics : Animals may only be treated with antibiotics when necessary for treatment of illnesses, provided they are not slaughtered within 45 days of last treatment.

Prohibited Use of Growth Hormones: The use of hormone treatments, including implants, to enhance growth is not permitted.

Prohibited Use of Genetically Modified Livestock: Animals produced through embryo transfer and those whose genetic material has been altered are not permitted.

Animal Health and Safety are Protected: Animal nutrition on the farm results in superior health as related to breeding success, weight gain, and freedom from illness. Policies are in effect for low-stress handling, preventative health measures, and regular maintenance and repair of facilities so as to prevent injury.

Feed is Pasture Based and Waste Management Systems Employed: Cattle receive majority of nutritional intake through grazing activity and animal movement is directed based on cattleâĂŹs natural action and reaction to the situation. Manure resources are used to close the nutrient cycle on the farm, but only to the extent that overall nutrient levels are adequate and not excessive. Excess manure, if any, is put to good use off farm.

Pests are Controlled Using Preventative Measures, Cultural and Nutritional Controls: Preventative measures and/or cultural controls such as movement of cattle, sanitation, and composting are used to reduce or eliminate the need for insecticides and miticides. Animals are free to choose and move to habitats that are most comfortable such as shady areas, windy spots

REFERENCES

REFERENCES

Batte, M.T. 2010. "Consumer-driven changes in food marketing channels: organics and sustainable food systems in the United States: Discussion."American Journal of Agricultural Economics 93 (2): 604-605.

Bond, C.A., D. Thilmany, and J. Kelling-Bond. 2008. "Understanding consumer interest in product and process based attributes for fresh produce." Agribusiness 24: 231-252.

Callens, I. and D. Tyteca. 1999. "Towards indicators of sustainable development for firms: A productive efficiency perspective." Ecological Economics 28: 41-53.

Casini, L. and A. M. Corsi. 2009. "Consumer preferences of wine in Italy applying best-worst scaling." International Journal of Wine Business Research 21(1): 64-78.

Clonan, A., Holdsworth, M., Swift, J., and Wilson, P., 2010. "UK consumers priorities for sustainable food purchases." Available at http://purl.umn.edu/91948.

Cohen, E. 2009. "Applying best-worst scaling to wine marketing." International Journal of Wine Business Research 21: 8-23.

Flynn, A., and J.J. Louviere. 1992. "Determining the appropriate response to evidence of public concern: the case of food safety." Journal of Public Policy and Marketing 11:12-25.

Flynn, T.N., J.J. Louviere, T.J. Peters, and J. Coast. 2007. "Best-worst scaling: What it can do for health care research and how to do it." Journal of Health Economics 26:171-189.

Lusk, J.L. and B. Briggeman. 2009. "Food Values." American Journal of Agricultural Economics 91: 1-13.

Lusk, J. and N. Parker. 2009. "Consumer preferences for amount and type of fat in ground beef." Journal of Agricultural and Applied Economics 41(1):75-90.

Marley, A.A.J., and J.J. Louviere. 2005. "Some probabilistic models of best, worst, and best-worst choices." Journal of Mathematical Psychology 49: 464-480.

Mueller, S., and C.Rungie. 2009. "Is there more information in best-worst choice data? Using attitude heterogeneity structure to identify consumer segments." International Journal of Wine Business Research 21: 21-40.

Onozaka, Y. and D. Thilmany McFadden, 2010. "Defining sustainable food market segments: do motivations and values vary by shopping locale?" American Journal of Agricultural Economics 93(2): 583-589.

Onozaka, Y. and D. Thilmany McFadden. 2011. "Does local labeled complement or compete with other sustainable labels? A conjoint analysis of direct and joint values for fresh produce claims." American Journal of Agricultural Economics 93 (3): 693-706.

Saunders, C., M. Guenther, W. Kaye-Blake, S. Miller, and P. Tait. 2010. "Consumer attitudes towards sustainability attributes on food labels." New Zealand Agricultural and Resource Economics Society Inc. Available at http://purl.umn.edu/96944.

Scarpa, R., S. Notaro, J. Louviere, and R. Raffaelli. "Exploring Scale Effects of Best/Worst Rank Ordered Choice Data to Estimate Benefits of Tourism in Alpine Grazing Commons." American Journal of Agricultural Economics 93(3): 813-828.

L L.Thurstone . "The Method of Paired Comparisons for Social Values, "Journal of Abnormal and Social Psychology, 21, (1927): 384-400.

Umberger, Wendy J., Stringer, Randy and Mueller, Simone C., 2010. "Using Best-Worst Scaling to Determine Market Channel Choice by Small Farmers in Indonesia" 2010 Annual Meeting, July 25-27, 2010, Denver, Colorado 90853, Agricultural and Applied Economics Association.

USDA NIFA. http://www.nal.usda.gov/afsic/pubs/terms/srb9902.shtml#toc2

Chapter 3: Consumer Preferences for Sustainably Produced Steak and Apples

Introduction

As suggested by our results in Chapter 2, sustainability, as a concept, remains elusive to many. The word "sustainable" has been attributed to everything from the foods we eat, to the structures we build, to the way we raise our families and run our businesses. While there is certainly a breadth of literature on sustainability, exploring how it is measured and managed, there are surprisingly few studies examining sustainability attributes in the context of food production. Two studies, to our knowledge, address consumer attitudes and preferences towards sustainability attributes of food. First, Saunders et al (2010) use a discrete choice model to investigate consumer decisions to purchase sustainably labeled foods, after displaying information about reduced carbon emissions. Alternatively, Clonan et al (2010) assess attitudes towards sustainably produced food products through a questionnaire about previous purchases and a Likert scale ranking of preferences. Both studies were used to inform the choice of sustainable farming attributes addressed in our survey. This work addresses a gap in the literature by evaluating preferences for sustainably labeled food products within the context of a choice experiment with varying degrees of information on sustainable agricultural systems.

Alternative agricultural production systems generate value-added food attributes that are used for product differentiation and provide consumers with a means by which to control their ecological footprint through their consumption choices. "Sustainably produced" food labels have rapidly grown in popularity over the past decade (Batte 2011). However, because there is no government agency overseeing certification of these production methods, evidence suggests that consumers are generally confused about the production attributes that may or may not be present in a sustainable food system. This paper analyzes data from a hypothetical choice experiment to better understand consumer purchasing behavior when faced with competing food production attributes such as "organic" and "local".

Choice experiments have a long history in the measurement of passive use values in environmental economics. Specifically, passive use value refers to the economic value derived from a change in environmental quality that may not be directly observable in market behavior (Adamowicz 1998). That is, changes in food production attributes create an additional dimension of consumer utility, which may be traded for other quality indicators, such as price and certification. In real markets, consumers are faced with consumption choices over bundles of attributes that can be modeled in a stated preference framework.

Marketing, transportation, and psychology literature led the field for some time in the adaptation of choice experiments. The method arose from conjoint analysis, but differs in the choice task to be completed. Unlike conjoint analysis, which utilizes scaled ranking or rating systems, choice experiments more closely mirror the attribute bundles of competing alternatives that are found in real markets. Choice experiments are compatible with random utility theory and are thus useful for determining the share of preference a given attribute has in a particular market. Therefore, hypothetical choice experiments provide a richer description of the attribute trade-offs that consumers are willing to make, than more traditionally used contingent valuation methods.

In the past, choice experiments have been used successfully by economists to measure the effect of environmental improvements and the value of quality differentiation (Boxall and Adamowitz (1998)). Other economists have also employed choice experiments to value food attributes such as local, organic and natural as well as more intrinsic values in the realm of food nutrition and safety such as traceability, animal welfare and genetic modification. These studies generally show that, on average, consumers are willing to pay positive price premiums for food produced outside of the conventional agricultural model. For example, Lusk, Norwood and Pruitt (2006) find positive price premiums associated with a ban on sub-therapeutic antibiotics in the pork production industry and Liljenstolpe(2010) finds that consumers indicate preference for food safety and animal welfare dimensions of value-added pork attributes. Additionally, Onazaka et al (2008) find supporting evidence of higher price

premiums on organically certified and locally grown produce. Most closely related to this work, Clonan et al (2010) find that stated purchasing behavior demonstrates that free range and local products have higher market preference over other sustainability attributes.

We seek to estimate preferences for "sustainably produced" food products in a discrete choice modeling framework and to determine how preferences may be affected by varying degrees of information about sustainable agricultural systems. If there is a prevalence of confusion about the attributes implied by sustainable labeling schemes, it is important to know if providing information about specific sustainable production attributes can lessen this uncertainty. Moreover, if information is successful in decreasing uncertainty over choice, it could be used towards better market differentiation between sustainable and organic principles. Additionally, the willingness to pay measures estimated in this paper provide insight into the value trade-offs perceived between current eco-labeling schemes, and the potential for differentiating "sustainably produced" products from their "organic" and "local" counterparts.

Research Methodology

Discrete Choice Experiments

In a discrete choice modeling framework, all respondents are assumed to be utility maximizers, facing a choice among competing alternatives that return different levels of utility. The analyst cannot directly observe respondent utility, but can observe attributes about the competing alternatives. Specifically, define a random utility function (U_{ij}) as the ith consumer's utility of choosing option j. Then,

$$U_{ij} = V_{ij} + \epsilon_{ij} \tag{5}$$

where V_{ij} is the deterministic component and ϵ_{ij} is the stochastic error (Adamowicz et al (1998)). Thus the probability that consumer i chooses alternative j is given by,

$$Prob(V_{ij} + \epsilon_{ij} \geq V_{ik} + \epsilon_{ik}) \; ; \forall k \in C_i \tag{6}$$

where C_i is the choice set of alternative bundles faced by consumer i. This choice experiment was constructed from alternative bundles of three attributes (label, certification, and price) with varying levels.

Alternative Specific Design

An alternative specific design was utilized to capture labeling scheme tradeoffs. In a more commonly used general design, the respondent is faced with a choice of the status quo against one or more competing alternatives in which all attributes are allowed to vary across all levels. Whereas, in an alternative specific design, each choice scenario presents each of the level alternatives of a given attribute (Label) while allowing the remaining attributes to vary across all levels. The respondents of our survey were faced with a number of choice scenarios, and asked to choose between the product labeled sustainable, the product labeled organic, the product labeled local, and the typical unlabeled product. The first three labeled products varied in certification and all products varied in price in each choice scenario. Additionally, each choice scenario presented the option of not buying any of the products presented. The inclusion of opt-out more realistically mirrors true market opportunities. By allowing respondents to opt-out of each choice scenario, we remove the assumption of market participation which will reduce potentially "forced" choices.

The label attribute took four levels; sustainable, organic, local, or typical. Each label was defined to describe the type of agricultural system within which the food was produced. The certification attribute took one of three levels; USDA, Private Third Party or Self. The "typical" product was reported to have no certification. All certifications referred to verification of all processes used in production and claims made by an accompanying labeling scheme. Private third party certification was verified by an independent entity unrelated to the farm of origin or retailer of the product. "Self" certification referred to labeling claims

made by the farmer producing the food. The price attribute took one of three discrete levels: (0.99/lb, 1.49/lb, or 1.99/lb) for apples and (5.99/lb, 8.99/lb, or 11.99/lb) for ribeye steak, as determined by the market at the time of the survey design. The following table outlines the choice experiment attributes and levels used in our design. For full descriptions and definitions please refer to the Appendix.

Table 18: Attributes and Levels Used in the Choice Experiment

Attribute	Levels
Label	Sustainable Organic Local Typical
Certification	USDA Private Third Party Self
Product Specific Prices	
Price (Apple)	0.99/lb 1.49/lb 1.99/lb
Price (Steak)	5.99/lb 8.99/lb 11.99/lb

Data Collection

The data used for this analysis were gathered in our 2010 national web-based survey of 1002 households, described in Chapter 2. Two survey versions were developed; one for apples and one for steak (beef). All respondents in the sample were randomly assigned to complete only one product version of the survey. The choice experiment portion of the survey was preceded by questions about perceived importance of varying sustainable farming attributes using a best-worst framework, analyzed in the previous Chapter, as well as several Likert scale ranking tasks assessing previous knowledge of agricultural production system characteristics and food consumption history. This combination of respondent information provides our analysis with a comprehensive assessment of the perceptions and preferences related to sustainability attributes. Immediately prior to the hypothetical choice experiment each respondent was provided with one of two information treatments about sustainable agricultural practices. A summary of survey participant demographic statistics is repeated here for reference.

Table 19: Demographic Variables and Summary Statistics of Survey Participants

Variable	Variable Description	Apple	Steak
Gender	1= Male; 2 = Female	1.476	1.516
	Total Participants	500	502
Age	Average Age in Years	51.48	50.92
Adults	Number of Adults in Household	2.062	2.048
Children	Number of Children in Household	0.48	0.51
Meals	Number of Meals/ week with Product	6.64	3.19
Shop	% of Total Shopping at Location		
	Grocery Store	81.97	81.16
	Health Food Store	8.49	9.20
	Food Co-op	2.05	1.90
	Convenience Store	3.0	4.42
	Farmers Market	3.72	5.24
	Butcher	4.38	2.92
Assistance	1 = On Food Assistance; 2 = Otherwise		
	1 = On Food Assistance	13.2 %	11.95 %
	2 = Not on Food Assistance	86.8 %	88.05 %
Education	Highest Level Completed		
	1 = Did not graduate from high school	2.2 %	2.59 %
	2 = Graduated from high school, no college	17.8 %	18.12 %
	3 = Attended college, no degree earned	28.8 %	33.67 %
	4 = Attended college, associates or trade degree earned	15.4%	12.55 %
	5 = Attended college, Bachelors degree earned	24.6 %	22.52 %
	6 = Graduate or advanced degree earned	11.2 %	10.56 %
Household Income	Range of Pre-tax Income		
	1 = Less than $20,000	20.8%	19.12 %
	2 = $20,000 - $39,999	26.4%	28.88 %
	3 = $40,000 - $59,999	18.4 %	21.51 %
	4 = $60,000 - $79,999	17.6 %	11.95 %
	5 = $80,000 - $99,999	8.2 %	7.57 %
	6 = $100,000 - $119,000	3.8 %	4.98 %
	7 = $120,000 - $139,999	1.8%	2.39 %
	8 = $140,000 - $159,999	1.4 %	1.20 %
	9 = More than $160,000	1.6%	2.39 %

Information Treatments

Half of the respondents were randomly assigned to receive general information about sustainable agriculture from the USDA website. This information treatment outlined general principles of sustainability such as, "resource conserving", "socially supportive", and "economically viable". Alternatively, the remaining half of respondents received information about sustainable agricultural practices that are components of a sustainable certification scheme from the aforementioned third party certifier, Food Alliance. This information treatment provided eight detailed standards of sustainable agriculture such as energy conservation and waste recycling, reduced use of chemical inputs, and fair and ethical treatment of workers and livestock. Please see the Appendix for full information treatments.

Following the information treatment, a brief cheap talk script was included to mitigate the problems associated with hypothetical bias. The effectiveness of cheap talk scripts has been repeatedly confirmed in the literature. Notably, Aadland and Caplan (2006) suggest neutral scripts that avoid assumptions about positive bias, as different subsamples may react differently dependent on factors such as market familiarity. Tonsor and Shupp (2011) utilize a large national survey and split-sample experimental design, finding that cheap talk scripts may not only influence the level of willingness to pay estimated for representative consumers, but also, in general, produce more reliable estimates.

To maintain orthogonality and independence across our choice experiment a main effects orthogonal experimental design was employed. In a main effects orthogonal design, a subset of the full factorial design is selected such that all linearly additive utility terms are identifiable. Specifically, the ORTHOPLAN procedure in SPSS Conjoint identified an orthogonal design based on two attributes, each with three levels, yielding a choice set of eighteen alternatives. The design utilized in this study is balanced, that is, each level of the non-alternative specific attributes (price and certification) occurs with equal frequency across the entire choice set. Therefore, each attribute has equivalent statistical power in explaining preference. For the sake of brevity, and to lower the complexity of the overall

task, the respondents were randomly assigned to answer one of two blocks of nine questions. The order of the alternatives was randomized for participants to mitigate any ordering bias. An example of a question faced by all apple survey respondents follows:

Survey Question Example

Which one of the following apple displays listed below would you choose to purchase from?

Table 20: Choice Experiment Survey Question Example

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Private 3rd Party	USDA	Self		
Price	\$1.49 /lb	\$1.99 /lb	\$0.99 /lb	\$1.49 /lb	
	☐	☐	☐	☐	☐

Analysis

The choice experiment structure allowed each respondent to select from four versions of the food product: (1) sustainably labeled, (2) organically labeled, (3) locally labeled, (4) unlabeled, or to opt-out of the purchase. This data is analyzed in a random utility framework. In each scenario, the choice of an alternative represents a discrete choice from a set of competing alternatives. Each alternative in a given choice scenario can be represented with a utility function that contains a deterministic and a stochastic component. An individual will choose a given alternative only if the utility representing that alternative is greater than all other options in the choice scenario. If the stochastic component ϵ_{ij} is distributed i.i.d. extreme value then Adamowicz et al (1998), Boxall and Adamowicz (2002) and Lusk et al (2003) have shown that the probability of consumer i choosing alternative j is equal to,

$$Prob(j) = \frac{e^{sV_{ij}}}{\sum_{k \in C} e^{sV_{ik}}} \tag{7}$$

where s is a scale parameter. In a single sample the scale parameter cannot be directly identified and is thus assumed to be 1, according to Lusk et al (2003). However, if data is collected from more than one sample, the relative scale parameter can be calculated and accounts for the difference in the variation of unobserved effects on error variance heterogeneity. To test whether data from the two information treatments (USDA versus Food Alliance) can be pooled, we estimate the relative scale parameter, across the two data sets to control for potential differences in variance. A likelihood ratio test of parameter equality, after controlling for scale, is undertaken to determine if pooling is appropriate for the following parametric analysis. The parameters contributing to the probability of choice are then first estimated in a multinomial logit model (MNL).

For choice problems involving three or more alternatives, the multinomial technique is most often employed. In the multinomial framework, the data on a dependent variable can fall into one of several mutually exclusive categories. However, analysis is not necessarily straightforward, as there are many different models for the probabilities of the multinomial distribution. Model selection depends on whether or not some of the individual specific regressors vary across alternatives and whether the categories are ordered or unordered (Cameron and Trivedi 2009). For the purposes of this analysis, we employ and compare the fit of three models: the multinomial logit, the conditional logit, and the random parameters logit.

In each of these models, the left hand side variable represents choice and is coded as a dummy variable, taking the value of 1 indicating the alternative was chosen and 0 otherwise. Therefore each choice scenario yields five data observations for every respondent, corresponding to the four competing product alternatives plus opt-out. The right hand side variables reflect the attribute levels of each product that is available for purchase. To avoid the assumption of equal marginal trade-offs, certification enters the model as an effects-coded variable. For the purposes of this study, we allowed USDA to act as the base-case since consumers are likely most familiar with USDA certification and labeling. Thus, two

variables were included for certification, for self and private third party, respectively. Each certification variable took a value of -1 to indicate USDA certification, 1 to indicate certification accordingly and 0 otherwise. The remaining right hand side variables are alternative specific dummy variables corresponding to each of the four label attributes. Price was coded as a continuous variable. Therefore, the logistic regression to be estimated for participant i for each alternative j is as follows,

$$\begin{aligned} Choice_{ij} &= \beta Price_{ij} + \delta_1 CertP_{ij} + \delta_2 CertS_{ij} \\ &+ \gamma_1 LabelS_{ij} + \gamma_2 LabelO_{ij} + \gamma_3 LabelL_{ij} + \gamma_4 LabelT_{ij} \end{aligned} \tag{8}$$

The parameter on price (β) approximates mean marginal utility of income and the parameters on each certification variable (δ_1 and δ_2) indicate the marginal (dis)utility associated with a change from USDA certification to Private Third Party certification or Self certification, respectively. The parameters on each label variable (γ_1, γ_2, γ_3, and γ_4) indicate the marginal utility gained from the labeling claim (or lack thereof) on each product relative to opt-out.

Thus, average willingness to pay for each label attribute, ceteris paribus, can be calculated as,

$$WTP(Label_i) = \frac{\gamma_i}{\beta} \tag{9}$$

and the price premium that can be captured on average for each labeling scheme, relative to the typical (unlabeled) product is,

$$Premium(Label_i) = \frac{(\gamma_i - \gamma_4)}{\beta} \tag{10}$$

The parameters of multinomial models are not directly interpretable in general. The es-

timated parameters will be transformed into marginal effects measures (WTP) to reflect the increased (or decreased) probability of choice accompanying a one unit change in the attribute level of each variable. Thus, it is anticipated that price and certification variables will exhibit negative coefficient estimates, as increased price and certification other than USDA should provide disutility to the consumer under our model assumptions.

Multinomial Logit

The multinomial logit model is best for datasets that contain only case-specific variables. It is the simplest model computationally and yields parameter results that are the easiest to interpret. To ensure model identification, the coefficient on one of the explanatory variables is set to zero and all other coefficients interpreted with respect to that category. For example, in our specification we have identified the base category as the no purchase option to make all coefficients relative to market participation. The multinomial logit requires data to be in wide-form, as explained in Chapter 2.

Conditional Logit

A closely related technique, the conditional logit, is a model also suited for behavioral modeling of polychotomous choice situations. Developed by McFadden in 1973, the conditional logit model is an arguably preferable technique for estimating the class of models in which choice among alternatives is a function of the characteristics of the alternatives rather than characteristics of the individual making the choice, such as quality measures for the alternatives, which is the case here. The conditional logit model requires that the data be in long form, as discussed in the previous Chapter. All case-specific variables will appear as a single variable that takes the same value for the five alternative outcomes. The two models are similar and perfectly symmetric in their statistical specification. The only difference is that the explanatory variables are open to assume different values in each alternative in the CL model, although the impact of a one unit change in an explanatory variable is assumed to be

constant across alternatives. It should be noted that multinomial logit and conditional logit models assume that all respondents share the same coefficients for a given attribute, an assumption of consumer homogeneity across preferences. This assumption is likely unrealistic if there are expectations that consumer preferences are in fact heterogeneous.

Random Parameters Logit

There are two primary strategies for addressing consumer heterogeneity in random utility models that differ in their assumptions about the distribution of preferences (Sagebiel (2011)). The first is to employ a random parameters logit (RPL) estimation on the same data set. The RPL accommodates heterogeneity in the data as a continuous function of the parameters and is free of the independence of irrelevant alternatives (IIA) assumption. The RPL is a more flexible method for modeling heterogeneity, compared to the latent class model employed in the previous Chapter, because it can induce any behavioral assumption in terms of preference distribution (Sagebiel). The conditional logit parameters are assumed fixed, taking the same value for all respondents, whereas the RPL is characterized by randomness in parameters. The unconditional RPL, used here, measures choice probability as a weighted average of all possible coefficients for the attribute parameters. The maximum likelihood estimates are achieved through simulation.

In contrast, the conditional logit specification can model heterogeneity by incorporating interaction terms of case-specific demographic variables or perception rankings. The drawback with this method is that the interaction terms introduced are motivated by the researcher's choice, and therefore are more exploratory in nature and not included in this analysis. The conditional and random parameters logit specifications are compared with the multinomial logit model by means of a likelihood ratio test to determine the best fit to the data. Furthermore, coefficient choice probabilities and willingness to pay estimates will be contrasted.

Results and Discussion

Information Treatment Tests

To specify our final model, we first estimated two separate multinomial logit models corresponding to each of the information treatments. Then the pooled data was used to estimate the same model, where utility parameters were constrained to be equal across information treatments. The test for parameter equality yields a test statistic that is calculated as follows (Lusk (2006)),

$$\lambda = -2(LL_p - (LL_u + LL_f)) \tag{11}$$

where LL_p is the log likelihood value of the pooled model after controlling for scale, LL_u is the log likelihood value of the USDA information treatment model and LL_f is the log likelihood value of the Food Alliance information treatment model. The test statistic, λ is distributed χ^2 with $K(M-1)$ degrees of freedom, where K is the number of restrictions (seven) and M is the number of information treatments (two).

Table 21: H_0: Pooling Across Information Treatments: Apple Survey

Subsample Modeled	n	LL	p-Value
All Respondents	502	-11198.3910	
USDA Information Treatment	251	-5568.8387	
Food Alliance Information Treatment	251	-5623.3510	
H_0 : *Pooling across Information Treatments is Okay*			0.8807

Table 22: H_0: Pooling Across Information Treatments: Beef Survey

Subsample Modeled	n	LL	p-Value
All Respondents	500	-10517.9060	
USDA Information Treatment	248	-5320.7855	
Food Alliance Information Treatment	252	-5193.1726	
H_0 : *Pooling across Information Treatments is Okay*			0.3419

Comparing the two likelihood function values using the likelihood ratio test resulted with associated two-tailed p values of 0.8807252 for apples and 0.34187185 for steak, indicating that the null hypothesis of parameter equality across information treatments cannot be rejected at any standard level of significance.

The likelihood ratio test results imply that information had no statistically significant effect on choices made in the hypothetical choice experiment across labeling schemes. Consumers with more specific information about the production standards required for third party certification exhibited preferences similar to those that received more general information about sustainable principles from the USDA. An interesting extension of this analysis could have included a third subsample, randomly assigned to receive no information, as was done in Lusk, Norwood and Pruitt (2006). The insignificance of the information treatment may be due to competing preconceptions about sustainability attributes, or it may be the persistence of general confusion over how these sustainable principles are implemented in practice.

Tables 23 and 24 present our primary estimation results. The MNL results are as expected, as price coefficients are negative and significant, certification parameters are negative and significant (with the exception of CertS for the USDA Apple sample), and label parameters are positive and significant.

Table 23: Information Treatment Comparison: Apple Survey

Variable	Pooled	Treatment USDA	Treatment FA
Price	-2.140	-2.118	-2.161
	(0.537)*	(0.076)*	(0.076)*
CertP	-0.250	-0.206	-0.291
	(0.031)*	(0.043)*	(0.044)*
CertS	-0.076	-0.157	-0.058
	(0.030)**	(0.043)	(0.042)*
LabelS	1.803	1.820	1.784
	(0.079)*	(0.113)*	(0.113)*
LabelO	1.884	1.811	1.949
	(0.079)*	(0.113)*	(0.113)*
LabelL	2.071	1.984	2.152
	(0.080)*	(0.114)*	(0.113)*
LabelT	0.603	0.547	0.656
	(0.085)*	(0.121)*	(0.119)*

(*) and (**) significant at 0.01 and 0.05 levels, respectively

Table 24: Information Treatment Comparison: Beef Survey

Variable	Pooled	Treatment USDA	Treatment FA
Price	-0.379	-0.358	-0.402
	(0.010)*	(0.013)*	(0.014)*
CertP	-0.321	-0.305	-0.337
	(0.033)*	(0.047)*	(0.048)*
CertS	-0.191	-0.155	-0.228
	(0.032)*	(0.046)*	(0.047)
LabelS	1.853	1.685	2.029
	(0.085)*	(0.119)*	(0.122)*
LabelO	1.847	1.702	1.998
	(0.085)*	(0.119)*	(0.122)*
LabelL	1.866	1.694	2.048
	(0.085)*	(0.119)*	(0.123)*
LabelT	0.818	0.672	0.972
	(0.089)*	(0.124)*	(0.127)*

(*) and (**) significant at 0.01 and 0.05 levels, respectively

Pooled Model Results

Moving forward with the pooled sample, three models were run on the data: the multinomial logit, the conditional logit, and the random parameters logit. In general, the results indicate that on average, as determined by the relative magnitudes of the coefficient estimates, the preference ordering for labeling claims on apples are as follows in descending order: 1) Local 2) Organic 3) Sustainable and 4) Typical. All parameters on Label are positive and significant at the 1% level in all three models. For beef, the results indicate that on average, as determined by the relative magnitudes of the coefficient estimates, the preference ordering for labeling claims on ribeye steaks are such that local is the most preferred, followed by an equal preference across organic and sustainable labels, and the typical product least preferred. Similar to the apple results, all parameters on Label are positive and significant in all three models, again at the 1% level.

The negative and statistically significant estimates on private third party certification and self certification indicate that consumers receive disutility from either of these certification schemes relative to USDA certification, ceteris paribus. Accordingly, the relative magnitude of the two certification coefficients reveals that for both apple and beef samples, Self certified products are least favorable, holding all other quality attributes constant. Also, as expected the price coefficient is negative and significant for both apple and steak samples. The discrete response profiles of choice frequencies across the five alternatives and coefficients for all three models are reported in the following tables.

Table 25: Discrete Response Profile: Apple Survey

Choice	Frequency	Percent
Sustainable	1064	23.25 %
Organic	1123	24.86 %
Local	1267	28.04 %
Typical	415	9.19 %
No Purchase	649	14.36 %

Table 26: MNL, CL, and RPL Parameter Estimates: Apple Choice Experiment

Variable	MNL	CL	RPL
Price	-2.14*	-2.13*	-2.35*
CertP	-0.25*	-0.25*	-0.30*
CertS	-0.07**	-0.05	-0.11*
LabelS	1.80*	3.56*	3.61*
LabelO	1.88*	3.64*	3.75*
LabelL	2.07*	3.82*	3.93*
LabelT	0.60*	2.39*	2.73*
Log Likelihood	-11198.39	-8982.39	-5963.00

(*) and (**) significant at 0.01 and 0.05 levels, respectively

Table 27: WTP Estimates for Apple Labeling Scheme Attributes

Welfare Measure	Model Used MNL	Model Used CL	Model Used RPL
WTP LabelS	$0.84/lb	$1.67/lb	$1.54/lb
WTP LabelO	$0.88/lb	$1.71/lb	$1.60/lb
WTP LabelL	$0.97/lb	$1.79/lb	$1.67/lb
WTP LabelT	$0.28/lb	$1.12/lb	$1.16/lb
Premium(LabelS)	$0.56/lb	$0.55/lb	$0.37lb
Premium(LabelO)	$0.60/lb	$0.59/lb	$0.43/lb
Premium(LabelL)	$0.69/lb	$0.67/lb	$0.51/lb

Using Log Likelihood as a measure of fit, we see that the Random Parameters Logit out-performs the multinomial and conditional logit models for apple responses. This result indicates significant evidence of heterogeneity in consumer preferences for our sample. The RPL model yields absolute WTP estimates in between those of the MNL and CL models, but lower price premiums overall. Applying a simple welfare measurement to the parameter estimates, we find that consumers are willing to pay $\$1.67/lb$ on average for locally produced apples, $\$1.60/lb$ on average for organic apples, $\$1.54/lb$ on average for sustainably labeled apples, and $\$1.16/lb$ on average for typical, unlabeled apples. These results yield price premiums of $\$0.51/lb$ on local apples, $\$0.43/lb$ on organic apples and $\$0.37/lb$ on sustainably produced apples, relative to their typical, unlabeled counterpart.

Table 28: Discrete Response Profile: Steak Survey

Choice	Frequency	Percent
Sustainable	971	21.58 %
Organic	967	21.49 %
Local	979	21.76 %
Typical	424	9.42 %
No Purchase	1159	25.76 %

Table 29: MNL, CL, and RPL Parameter Estimates: Steak Choice Experiment

Variable	MNL	CL	RPL
Price	-0.38*	-0.38*	-0.38*
CertP	-0.32*	-0.32*	-0.34*
CertS	-0.19*	-0.16*	-0.23*
LabelS	1.85*	2.89*	2.97*
LabelO	1.85*	2.88*	2.99*
LabelL	1.87*	2.91*	2.99*
LabelT	0.82*	1.88*	2.12*
Log Likelihood	-10517.91	-9027.11	-5985.00

(*) and (**) significant at 0.10 and 0.05 levels, respectively

Table 30: WTP Estimates for Steak Labeling Scheme Attributes

Welfare Measure	Model Used MNL	Model Used CL	Model Used RPL
WTP LabelS	$5.00/lb	$7.61/lb	$7.81/lb
WTP LabelO	$5.00/lb	$7.58/lb	$7.86/lb
WTP LabelL	$5.02/lb	$7.66/lb	$7.86/lb
WTP LabelT	$2.22/lb	$4.95/lb	$5.58/lb
Premium(LabelS)	$2.78/lb	$2.66/lb	$2.24lb
Premium(LabelO)	$2.78/lb	$2.63/lb	$2.29/lb
Premium(LabelL)	$2.81/lb	$2.71/lb	$2.29/lb

Using Log Likelihood as a measure of fit, we see again that the Random Parameters Logit out-performs the multinomial and conditional logit models for the steak responses. Applying the same welfare measurement to the steak participant subsample, we find that consumers are willing to pay $\$7.86/lb$ on average for locally or organically produced steak, $\$7.81/lb$ for sustainably produced steak, and $\$5.58/lb$ for typical unlabeled steak. These results yield price premiums of $\$2.29/lb$ on local and organic steaks, and $\$2.24/lb$ for sustainable steaks relative to their typical, unlabeled counterpart.

All welfare measurements should be interpreted relative to opt-out. That is, the willingness to pay on each label attribute is the amount of money the average consumer is willing to pay for each labeled product after choosing to participate in the market, holding all else constant and assuming no uncertainty regarding choice. Accounting for heterogeneity, by employing the random parameters logit to model this data, WTP estimates fall in the range of true market value for these products at 2010 price levels.

The positive willingness to pay estimates on all labels imply that all of the products individually are preferred to not buying any. On average, consumers are willing to pay the most for locally produced apples. These results imply that consumers prefer locally produced and organically produced steaks equally. Following the preference for locally grown and produced apples and steak, organic and sustainably labeled products have the next highest willingness to pay, for apples and steak respectively. However, in both cases the willingness to pay welfare measurements are close in magnitude, which could easily lead to a reversal of the preference ordering for these two label attributes if this experiment was replicated with another population.

WTP comparison

The WTP estimates reported above are calculated from the RPL mean valuations of model covariates. The assumptions inherent in this model ignore the distribution of preferences around the mean of random parameters. We relax the strength of this assumption by fol-

lowing the simulation techniques prescribed in Hensher et al (2006) and similarly utilized in Tonsor et al (2009). Specifically, we consider the entire distribution of WTP and the statistical variability in parameter estimates to capture heterogeneous preferences and to empirically test the null hypothesis of equal WTP preferences across (alternative specific) labels.

95% confidence intervals of WTP are derived from generating a distribution of 999 WTP estimates using a parametric bootstrapping technique proposed by Krinsky and Robb (1986). The parametric bootstrapping method essentially draws 999 observations from multivariate normal distributions around the coefficients from our logit model by resampling from the original sample with replacement. For tests at the α level or for $100(1-\alpha)\%$ confidence there are reasons for choosing the number of draws, B, such that $\alpha(B+1)$ is an integer. Therefore, we follow the advice of Cameron and Trivedi (2010) and use $B = 999$ for confidence intervals and hypothesis tests when $\alpha = 0.05$. Mean WTP estimates and 95% confidence intervals are identified incorporating both statistical and preference variability. An initial evaluation of overlapping 95 % confidence intervals yields a general assessment of WTP differences.

Table 31: Hypothesis Test of Equal WTP Across Credence Labels: Apple Survey

Label	Average	95% interval
WTP LabelS	$1.61/lb	($1.45/lb, $1.79/lb)
WTP LabelO	$1.73/lb	($1.57/lb, $1.93/lb)
WTP LabelL	$1.78/lb	($1.60/lb, $1.97/lb)
WTP LabelT	$1.14/lb	($1.00/lb, $1.30/lb)

Table 32: Hypothesis Test of Equal WTP Across Credence Labels: Steak Survey

Label	Average	95% interval
WTP LabelS	$7.75/lb	($6.87/lb, $8.74/lb)
WTP LabelO	$7.75/lb	($6.87/lb, $8.74/lb)
WTP LabelL	$7.83/lb	($6.92/lb, $8.82/lb)
WTP LabelT	$4.97/lb	($4.26/lb, $5.80/lb)

The bootstrapped estimates reveal that the 95% confidence intervals on all three credence labeling claims overlap for both apple and steak survey respondents. This result indicates that consumers do not differentiate between sustainable, organic, and local labels. However, because none of the three credence labeling attribute intervals overlap with the typical, unlabeled product interval we can conclude that each of the labels individually is successfully differentiated from the unlabeled product.

Conclusions

The objectives of this paper were to employ a stated preference approach, utilizing a choice experiment framework, for measuring value associated with quality changes in sustainable agricultural production practices, and to determine if providing information about sustainable agricultural practices affects willingness to pay estimates for sustainably labeled food products. Our results support previous studies of food attribute valuation (Bond et al (2008), Lusk and Briggeman (2009)) and illustrate the higher preference for locally grown and produced foods (Detoni and Tonsor (2009)). Moreover, our analysis provides evidence of weak market differentiation between sustainable, organic and locally labeled products, which may further exacerbate uncertainty regarding the attributes associated with these credence labeling schemes. Abram et al (2010) found similar results when evaluating perceptions of "all natural" claims against organic. Providing varying levels of information on sustainable agricultural practices did not yield significantly different preferences across respondents of our survey.

Our analysis provides significant evidence of consumer heterogeneity in our sample and demonstrates that positive price premiums can be captured by sustainably produced labeling claims, relative to similar unlabeled and conventionally produced food products. However, the price premiums calculated in our model reveal that there is a comparable tradeoff in quality associated between local, organic and sustainably labeled food products. Furthermore, detailed information about sustainable certification guidelines had no significant impact on choosing the sustainably labeled products. Thus, we conclude that, based on our results, consumer demand for sustainably produced food may not be distinctly differentiable from its local and organic counterparts. Overall, these findings suggest that profitable marketing opportunities may exist for firms interested in selling sustainably produced food products, however there needs to be considerable effort put into leveling information asymmetries about product quality if sustainable label claims are to be differentiated from the local food movement or more recognizable organic principles.

APPENDICES

USDA Information Treatment

Please consider the following information [provided by the United States Department of Agriculture]:

The word â*sustain* from the Latin sustinere (sus, from below and tenere, to hold), to keep in existence or maintain, implies long-term support or permanence. As it pertains to agriculture, sustainable describes farming systems that are capable of maintaining their productivity and usefulness to society indefinitely. Such systems must be resource-conserving, socially supportive, commercially competitive, and environmentally sound. Sustainable agriculture was addressed by Congress in the 1990 Farm Bill and the Food, Agriculture, Conservation, and Trade Act of 1990 (FACTA). Under that law, the term sustainable agriculture means an integrated system of plant and animal production practices having a site-specific application that will, over the long term:

1. Satisfy human food and fiber needs
2. Enhance environmental quality and the natural resource base upon which the agricultural economy depends
3. Make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls
4. Sustain the economic viability of farm operations
5. Enhance the quality of life for farmers and society as a whole

Consumers are increasingly demanding on the ethical dimension of food quality. This relates to the process of production and trade and its broad impacts on society and the environment. It includes a wide range of social, environmental or cultural issues such as the treatment of workers, a fair return to producers, environmental impacts and animal welfare. Guidelines about what specific practices meet long-term environmental, economic and social goals and constitute sustainable agriculture is still under debate. However, a

handful of groups have attempted to develop standards and/or provide certification services based upon their own guidelines.

Food Alliance Information Treatment

Please consider the following information [provided by Food Alliance, an independent third party certifier of sustainably produced foods]:

The impacts of food production have become a mainstream concern. Expectations for traceability, transparency and accountability in agriculture and the food industry are increasing. Sustainable agriculture comprises the ability to produce safe, healthy, delicious, and affordable food to meet needs without degrading agricultural lands, the quality of life in our communities, or the resiliency of the broader ecosystems on which we all depend. Farms employing sustainability practices place important on issues including safe and fair working conditions, humane treatment of animals, and protection of the environment. Sustainable farms should be held to the following standards:

1. Protect, enhance, and conserve soil resources, water resources, and biodiversity

   Food production improves soil productivity, protects water quality and supply, and supports healthy native plant and animal communities.

2. Conserve energy, reduce and recycle waste

   Waste streams from food production are minimized while reuse, recycling, and composting of resources is maximized. Businesses invest in innovation and improvement to ensure efficient use and management of natural resources for energy and packaging, transport, and daily operations.

3. Reduce use of pesticides, and other toxic and hazardous materials

   Food businesses avoid use of chemicals that have adverse impacts on the health of ecosystems. Agriculture relies on a biologically based system of Integrated Pest Management. Materials used for sanitation, pest control, waste treatment, and infrastructural maintenance are chosen to reduce overall negative consequences.

4. Maintain transparent and sustainable chain of custody

   Farmers and food industry workers have secure and rewarding jobs that provide a sound livelihood. Throughout the entire supply chain, food is produced and handled in accordance with these Principle Values. Transparency is maintained independent standards, third-party audits and clear labeling.

5. Guarantee product integrity, no genetically engineered or artificial ingredients

   Foods are not produced using synthetic preservatives, artificial colors and flavors, genetically modified organisms (GMOs), or products derived from livestock treated with sub-therapeutic antibiotics or growth-promoting hormones.

6. Support safe and fair working conditions

   Employers respect workersâ rights and well-being, make safety a priority, maintain a professional workplace, and provide opportunities for training and advancement.

7. Ensure healthy, humane animal treatment

   Animals are treated with care and respect. Living conditions provide access to natural light, fresh air, fresh water, and a healthy diet, shelter from extremes of temperature, and adequate space and the opportunity to engage in natural behaviors and have social contact with other animals. Livestock producers minimize animal fear and stress during handling, transportation and slaughter.

8. Continually improve practices

   Food businesses are committed to continually improving management practices. Improvement goals are integrated into company culture, regularly monitored, and acknowledged when achieved. Food buyers are proactively engaged in the food system, and support companies that are transparent about their improvement goals and progress.

Choice Experiment Instructions, Definitions, and Cheap Talk Script

In the next section you will be presented with multiple different alternative packages of beef ribeye steak that could be available for purchase in a retail store where you typically shop. Besides the attributes listed below, each product possesses the same characteristics (e.g., similar color and freshness) and is produced in the U.S. Prices vary for each product and are all in $/lb. units. Please consider the following information to help you interpret alternative products.

Label: The package that contains the beef ribeye steak for your purchase may be labeled as follows:

- Sustainable: This beef was produced using sustainable practices.
- Organic: This beef was produced using organic practices.
- Local: This beef was produced for distribution and sale locally.
- Typical: This beef is not labeled to suggest it was produced using any of the criteria listed above.

Certification: The typical product has no certification label. Each labeled product can be certified in one of three ways:

- USDA: The processes used and all claims made by the product label have been verified by the USDA.
- Independent Third Party: The processes used and all claims made by the product label have been verified by a third party unrelated to the farm of origin or retailer.
- Self: The processes used and all claims made by the product label have been verified by the farmer producing the food.

The experience from previous similar surveys is that people often state a higher willingness to pay than what one is actually willing to pay for the good. For instance, a recent study asked people whether they would purchase a new food product similar to the one you are about to be asked about. This purchase was hypothetical (as it will be for you) in that no one actually had to pay money when they indicated a willingness to purchase. In the study, 80% of people said they would buy the new product, but when a grocery store actually stocked the product, only 43% of people actually bought the new product when they had to pay for it. This difference (43% vs. 80%) is what we refer to as hypothetical bias.

Accordingly, it is important that you make each of your upcoming selections like you would if you were actually facing these exact choices at a store; noting that buying a product means that you would have less money available for other purchases.

REFERENCES

REFERENCES

Aadland, D. and Caplan, A.J. 2006. Cheap talk reconsidered: new evidence from CVM. Journal of Economic Behavior and Organization 60, 562-578.

Abrams, Katie, Courtney Meyers, and Tracy Irani, 2010. Naturally Confused: Consumers' Perceptions of All-Natural and Organic Pork Products. Agriculture and Human Values 27: 365-374.

Adamowicz, W., Boxall, R., Williams, M., and Louviere, J., 1998. Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. American Journal of Agricultural Economics 80, 64-75.

Batte, M.T., 2010. Consumer driven changes in food marketing channels: organics and sustainable food systems in the united states: discussion. American Journal of Agricultural Economics 93, 604-605.

Bond, C.A., D. Thilmany, and J. Kelling-Bond, 2008. Understanding consumer interest in product and process based attributes for fresh produce. Agribusiness 24, 231-252.

Boxall, P. and Adamowicz, W., 2008. Understanding heterogeneous preferences in random utility models: A latent class approach. Environmental and Resource Economics 23, 421-446.

Clonan, A., Holdsworth, M., Swift, J., and Wilson, P., 2010. UK consumers priorities for sustainable food purchases. http://purl.umn.edu/91948

Darby, K., M. T. Batte, S. Ernst, and B. Roe, 2008. Decomposing Local: A conjoint analysis of locally produced foods. American Journal of Agricultural Economics 90, 476-486.

Detoni, D., and G.T. Tonsor, 2009. The direct and indirect effects of 'locally grown' on consumers attitudes towards agri-food products. Agricultural and Resource Economics Review 38, 384-396.

Liljenstolpe, C. 2010. Demand for value-added pork in Sweden: A latent class model approach. Agribusiness 27, 129-146.

Lusk, J.L. and Briggeman, B., 2009. Food values. American Journal of Agricultural Economics 91, 1-13.

Lusk, J.L., Norwood, F.B. and Pruitt, J.R. 2006. Consumer demand for a ban on antiobiotic drug use in pork production. American Journal of Agricultural Economics 88, 1015-1033.

Onozaka, Y. and D. Thilmany McFadden, 2010. Defining sustainable food market segments: do motivations and values vary by shopping locale? American Journal of Agricultural Economics; doi:10.1093/ajae/aaq152

Onozaka, Y. and Thilmany-McFadden, D. 2011. Does local labeling complement or compete with other sustainable labels? a conjoint analysis of direct and joint values for fresh produce claims. American Journal of Agricultural Economics 93, 693-706.

Sagebiel, J. 2011. Comparing the latent class model with the random parameters logit - a choice experiment analysis of highly heterogeneous electricity consumers in Hyderabad, India. Working paper. julian.sagebiel@hu-berlin.de

Saunder, C. et al. 2010. Consumer attitudes towards sustainability attributes on food labels. Presented paper at the New Zealand Agricultural and Resource Economics Society.

Shulz, L.L. and Tonsor, G.T., 2009. Cow-calf producer preferences for voluntary traceability systems. Journal of Agricultural Economics 61, 138-162.

Tonsor, G.T., Schroeder, T.C., Pennings, J.M. and Mintert, J. 2009. Consumer valuations of beef steak food safety enhancement in Canada, Japan, Mexico, and the United States. Canadian Journal of Agricultural Economics 57, 395-416.

Tonsor, G.T., and Shupp R. S. 2011. Cheap talk scripts and online choice experiments: looking beyond the mean. American Journal of Agricultural Economics 93, 1015-1031.

Tonsor, G., and Shupp, R., 2009. Valuations of 'sustainably produced' labels on beef, tomato, and apple products. Agricultural and Resource Economics Review 38, 371-383.

Chapter 4: Willingness to Pay for Sustainably Labeled Steak and Apples: An Application in Experimental Auctions

Introduction

A current trend in the US food system toward organic, sustainable, and local foods has created a need for more research on the valuation of such credence attributes. Of these trends, sales in the organic food sector have grown the fastest, encouraged by the strength of its broad promotion and marketing. Most of the organic industry's growth has occurred in the years since the establishment of the USDA's National Organic Program, in 2002, creating a system of standards for certifying and labeling organic products. However, unlike USDA-certified organic products, foods designated as sustainably produced carry no government-endorsed certification and bear no standardized label. The USDA National Institute of Food and Agriculture provides limited information on the purported sustainability of different agricultural practices and warns that, "guidelines about what specific practices meet long-term environmental, economic, and social goals and constitute sustainable agriculture are still under debate".

"Economic sustainability" is a term used to identify strategies for using available resources in a way that promotes efficiency and responsibility, with a goal of providing long-run benefits. Many interpret this to mean that a sustainable system supports and sustains local economies by investing in community businesses, including but not limited to, agriculture. Local foods have become more available following a dramatic increase in demand and succeeding a significant expansion of farmers markets across the country. The term "local" remains undefined, with individual interpretations abound. However, interest in locally grown foods continues to rise, matched, in response, by the inception of many state-sponsored marketing campaigns. Understanding these growing trends and how they interact is therefore worthy of further examination.

If sustainable agricultural systems possess similar environmental management practices

as organic and support local farms and the communities they feed, can the consumer differentiate between these competing claims? Comprehending consumer attitudes towards and preferences between these three trends, as well as determining WTP for these production attributes is of importance to producers and marketers in determining the effectiveness of their labeling schemes and marketing programs.

The primary goal of this research is to determine consumer preferences and WTP for foods labeled as sustainably produced when offered alongside their local and organic counterparts. The goal of this paper is complemented by three additional objectives: (1) compare hypothetical and non-hypothetical WTP estimates to determine the extent of hypothetical bias (2) use this combination of hypothetical and non-hypothetical choice data to understand the relationship between sustainable, organic, and local label preferences, and (3) develop a set of recommendations designed to guide industry-specific producers in the development and implementation of a local marketing plan.

To accomplish these objectives, data gathered from a series of experimental auctions held in the fall of 2010 will be utilized in several economic models of choice. The data obtained in experimental auction procedures is believed to be more representative of actual behavior as demand-revealing, non-hypothetical methods are used. The experimental auction participants form a subset of respondents to a nation-wide survey with a hypothetical choice experiment component. Collectively, the use of experimental auctions augmenting hypothetical choice experiment data yield a more detailed and reliable understanding of consumer behavior. The use of multiple sources of data will allow for a more robust understanding of how consumers define and value the Òsustainably producedÓ label.

Research Methodology

Experimental Auctions

A growing literature exists on the design, implementation and evaluation of experimental auctions with agricultural or food related applications. Due to the advantage of experimental

auction methods in isolating the effect of information provision, it has become an increasingly popular avenue for investigating the impact of labeling schemes on consumer WTP for food products. Applications of experimental auction procedures, used to evaluate information provision or labels, range most recently across health and nutritional information (Hellyer et al (2012) and Hobbs et al (2006)), country of origin labeling (Chern and Chang (2012)) , traceability and food safety (Lee et al (2011) and Nayga et al (2006)), genetic modification and biotechnology (Colson and Huffman (2011), Corrigan and Rousu (2011), Lusk et al (2004)), and other quality attributes such as Ògrass-fedÓ (Evans et al (2011) and Umberger (2002)), and Òhormone-freeÓ (Alfnes and Rickertsen (2003)).

A small number of known studies have used experimental auctions in the valuation of macro-level food system credence attributes. Rousu and Corrigan (2008) utilize experimental auctions to compare several alternative fair trade labels in order to determine the welfare loss from labels that inadequately inform consumers. More recently, Briggeman and Lusk (2011) use a model of inequality aversion and altruism paired with experimental auctions to investigate consumer WTP for organic foods and to better understand preferences for fairness and equity in the food system. To the authors' knowledge, our study is the first to employ experimental auctions to investigate consumer preferences and WTP for food products labeled with broadly interpreted sustainable production claims.

Data Collection

Experimental Design and Implementation

A series of five experiment sessions, held in October 2010 and attended by a total of seventy-six participants, informs this study. The participants were recruited by the lead author from three different grocery stores in the greater Lansing, Michigan area with the aid of a promotional poster. Grocery stores were targeted for recruitment for the purpose of attracting primary household food shoppers. Demographic statistics on the experiment participants are reported below.

All experiment sessions were held in a classroom space at Michigan State University's Pavilion for Agriculture and Livestock Education. At the time of their recruitment, each participant signed up for a specific experiment session by providing a name and phone number. In return, each recruit was given information about the session, including: date, time, compensation, directions to the facility, and a token green MSU pencil to incentivize follow-through. The recruited participant information sheet is available upon request. Participants were called one day before the session with a reminder.

Upon arrival at the Pavilion, participants were given a unique ID for anonymity purposes and then were instructed to begin an online survey at one of the mobile laboratory laptops. The survey accompanying the experiment was identical to our nationwide survey, disseminated earlier in the summer of 2010, eliciting information about perceptions and preferences towards sustainable farming practices and sustainably labeled foods. As each participant completed the online survey, they signed the informed consent and received a copy of the auction instructions. Auction instructions were read aloud, including an example bidding strategy from a proposed auction scenario, and then participants were invited to view each of the products that were to be auctioned.

Each experiment session consisted of a series of eight auctions corresponding to the eight food products available; a 1 lb bag of apples labeled as sustainable, organic, local, or unlabeled and a 0.5 lb rib-eye steak labeled as sustainable, organic, local, or unlabeled. The order of the auctions was randomized for each food product in every session. All products were roughly identical and sourced from the same farm that was certified organic, local to the participant population, and made farm-level claims of sustainable production. However, the products were stripped of their original labeling and given only one of the following labels for presentation to experiment participants: "Sustainable", "Organic", "Local" or "Typical". It should be noted that the new labels for experimental purposes made no false claims, but allowed the researchers to isolate the effect of the Òsustainably producedÓ credence label while controlling for other observed quality attributes such as color, consistency, or size.

This experiment utilized the second-price Vickrey auction with full bidding. This choice of auction mechanism is popular in the literature for its relative performance compared to alternatives. Additionally, this auction is easy to explain to participants and has been shown to adequately measure "on margin" bidding (Lusk and Shogren 2007). No reference prices were provided and no bids were revealed as auction rounds progressed. Participants were randomly assigned to a group of 6-8 people, against whom they would be bidding in each auction. The identity of group members was never revealed. Smaller groups were used to retain the incentive compatible nature of the auction, while keeping participants engaged. When the auctions were undertaken, each participant wrote down a WTP bid on a sheet of paper provided to them, labeled with the product auctioned in that round. Bid sheets were collected by one researcher and immediately recorded into an electronic spreadsheet. At the conclusion of all eight auctions, one auction for each food product was randomly selected as binding for purposes of payment.

Each participant was endowed with $25 to bid on the four 0.5 lb rib-eye steaks and $5 to bid on the four 1 lb bags of apples, knowing that only one round, for each product, would require binding payment. Participants were informed that this money was theirs and if they chose not to bid on any product, or did not win the binding auction they would go home with their full endowment. However, if they won the binding auction, they would go home with the product and their endowment less the second highest bid. After the binding auction was announced for each product, participants were called up, one at a time, to receive their endowments less any payments made for products won in the binding auctions, the corresponding food products if applicable, and a $25 gift card to the grocery store where they were recruited to compensate them for their time. Full disclosure of farm-of-origin information was provided to participants that took home food products.

In total, 76 participants were recruited from Lansing area grocery stores for the described series of experimental auctions. A comparative summary of demographic statistics can be found in the following table. In general, the participant population was more female,

younger, more frequently white, ate less beef in a given week, shopped at food co-ops and farmers' markets more often, were less likely to have dropped out of college before finishing an associate's degree, and had lower household income than the larger survey sample of 1002 households collected earlier in 2010.

Table 33: Summary Statistics of Survey and Experiment Samples

		National Survey Sample	Experiment Sample
Variable	Definition	Mean or Frequency	Mean or Frequency
Gender	1= Male	48 %	37 %
	2 = Female	52%	63 %
Age	Average Age in Years	51.48	31.43
Adults	Number of Adults in Household	2.06	2.35
Children	Number of Children in Household	0.48	0.41
Meals	Number of Meals/ week with Apples	3.19	3.98
Meals	Number of Meals/ week with Beef	6.64	4.32
Shop	% of Total Shopping at Location		
	Grocery Store	81.97	56.53
	Health Food Store	8.49	7.89
	Food Co-op	2.05	9.07
	Convenience Store	3.0	3.0
	Farmers Market	3.72	16.16
	Restaurant	4.38	5.46
Assistance	1 = On Food Assistance; 2 = Otherwise		
	1 = On Food Assistance	13.2 %	11.84 %
	2 = Not on Food Assistance	86.8 %	88.16 %
Education	Highest Level Completed		
	1 = Did not graduate from high school	2.59 %	3.95%
	2 = Graduated from high school, no college	18.12 %	18.42 %
	3 = Attended college, no degree earned	33.67 %	7.89 %
	4 = Attended college, associates or trade degree earned	12.55 %	40.79 %
	5 = Attended college, Bachelors degree earned	22.52 %	18.42 %
	6 = Graduate or advanced degree earned	10.56 %	10.53 %
Household Income	Range of Pre-tax Income		
	1 = Less than $20,000	19.12 %	51.32 %
	2 = $20,000 - $39,999	28.88 %	14.47 %
	3 = $40,000 - $59,999	21.51 %	14.47 %
	4 = $60,000 - $79,999	11.95 %	11.84 %
	5 = $80,000 - $99,999	7.57 %	1.32 %
	6 = $100,000 - $119,000	4.98 %	3.95 %
	9 = More than $120,000	5.99 %	2.63 %

Analysis

The main focus of the analysis presented here is to identify predictors of bid behavior by estimating several regression models with the auction data. For all regression models the dependent variable is the WTP bid. In keeping with the literature, this work employs a double-censored tobit specification, as bids obtained in the auction are censored on the left by zero and at the right by the endowment; \$5 for apples and \$25 for steaks. The regression of interest is specified as an unobserved latent variable, y_i^*,

$$y_i^* = x_i^{'}\beta + \epsilon_i \tag{12}$$

where ϵ_i $N(0, \sigma^2)$ and x_i is the $(KX1)$ vector of exogenous and fully observed regressors (Cameron and Trivedi 2010). The observed variable y_i is related to the latent variable y_i^* through the following rule:

$$y_i = \begin{cases} L: & y^* \leq L \\ y^*: & L < y^* < U \\ U: & y^* \geq U \end{cases}$$

where $[L, U]$ is the censored interval of observed values. The foregoing analysis uses maximum likelihood estimation under the assumptions that the regression error is homoskedastic and normally distributed. ML estimates of (β, σ^2) solve the first-order conditions from maximization of the log likelihood based on the density function of censored observations. These equations are nonlinear in parameters and therefore the solution uses an iterative algorithm.

A detailed summary of the data on the dependent variables, apple bids and steak bids respectively, provides insight into potential problems with using MLE on the parameters of the tobit model under the strict assumptions of homoskedasticity and normalcy of errors. Examining apple bids further reveals only moderate skewness of 1.24 and slight kurtosis of -0.14 after appropriate correction. Similar examination of steak bids yields a skewness of

1.14 and kurtosis of -0.52. These tests offer encouraging evidence of the consistency of the MLE estimation, as the dependent variable appears to be approximately normal. Therefore, we continue with results of the linear tobit without any transformation of the dependent variable.

Results and Discussion

Pooled Sample Comparison

An identical analysis to the one utilized in Chapter 3 was employed on the choice experiment survey data for the auction participant subsample. Before settling on the RPL model, the multinomial logit and conditional logit were considered. The MNL was deemed a worse fit by a log likelihood comparison. Overall, model fit of the CL was stronger and consistent with earlier results. However, the MNL and CL models dropped the local label variable because of collinearity with the organic label variable. Standard logit models do not take into account any unobserved factors that may persist across alternatives for a given decision maker. To take into account correlation in unobserved factors over alternatives, the RPL allows the preference coefficients to be different for each respondent. The collinearity issues in the MNL and CL models indicate that the choice of the local and organic products were highly correlated for this much smaller and much less diverse sample, perhaps partially due to inadequate sample size (n=40 for apples and n=36 for steak).

We first explored the difference between the choice experiment results from the national survey sample and the auction participant sample by testing the null hypothesis of parameter equality using a likelihood ratio test, as used in Chapter 3 to test the effect of the information treatment. Comparing the likelihood function values using the likelihood ratio test resulted with an associated two-tailed p value < 0.001 for both apple and steak surveys indicating that the null hypothesis of parameter equality across participant samples is rejected at any standard level of significance.

This result indicates that the auction participant sub-sample has significantly different (hypothetical) preferences for apples and steak than the larger national survey sample. Moreover, the auction participant sample results yielded a positive coefficient on price in both surveys. This can be interpreted to mean that higher prices led to a higher probability of choice across alternatives. It may be that this small sample views price as an indicator of

quality moreso than the national sample. Also, the auction participant sample included a large number of people recruited from a Health Foods grocery store that may be accustomed to paying higher prices for health differentiated foods.

Additionally, the coefficient on the typical label is negative in the apple survey of experiment participants, suggesting that this sample prefers to opt out of the market over purchasing the typical, unlabeled product. This result is supported by the fact that many of the participants were recruited directly from a Health Foods specialty grocery store. By their choice of shopping locale, those participants are essentially making the choice in reality that was indicated by their hypothetical purchasing data; they are opting out of the regular super market that carries the typical (unlabeled) products.

Table 34: Pooled Sample Comparison: RPL Estimates in Apple Survey

	Sample		
Variable	Pooled	Survey Only	Survey and Auction
Price	-1.90*	-2.13*	1.33**
CertP	-0.26*	-0.25*	-1.32*
CertS	0.00	-0.048	-0.264
LabelS	3.34*	3.55*	2.46*
LabelO	3.44*	3.63*	2.52*
LabelL	3.55*	3.82*	2.41*
LabelT	2.09*	2.39*	-1.68*

(*) and (**) significant at 0.01 and 0.05 levels, respectively

Table 35: Hypothesis Testing of Pooling Across Participant Sample: Apple Survey

Subsample Modeled	n	LL	p-Value
Pooled Sample	542	-9848.4917	
Survey Only Sample	502	-8982.3905	
Auction Participant Sample	40	-676.7408	
H_0 : *Pooling across Participant Samples is Okay*			< 0.001

Table 36: Pooled Sample Comparison: RPL Estimates in Beef Survey

	Sample		
Variable	Pooled	Survey Only	Survey and Auction
Price	-1.90*	-0.37*	1.29**
CertP	-0.26*	-0.32*	-5.99*
CertS	0.00	-0.16*	3.18*
LabelS	3.34*	2.88*	9.39*
LabelO	3.44*	2.88*	9.41*
LabelL	3.55*	2.90*	9.55*
LabelT	2.09*	1.88*	0.99**

(*) and (**) significant at 0.01 and 0.05 levels, respectively

Table 37: Hypothesis Testing of Pooling Across Participant Sample: Beef Survey

Subsample Modeled	n	LL	p-Value
Pooled Sample	536	-9743.3977	
Survey Only Sample	500	-9027.1083	
Auction Participant Sample	36	-643.1427	
H_0 : *Pooling across Participant Samples is Okay*			< 0.001

Raw Data Comparison

Eight participants self-identified as vegetarian and bid $0.00 in all four rounds of steak auctions and therefore were removed from the steak data analysis that follows. Vegetarians made up 13.16% of this sample population. According to a 2006 study, employed by the Vegetarian Resource Group, in a national poll only 6.7% of the research population identified as vegetarian or never ate meat, making our sample twice as vegetarian as the purported United States population. It may be hypothesized that the nature of the self-selected participant group may lead vegetarians, or other people with restrictive diets, to be more interested in food related studies. Furthermore, two participants were deemed "unengaged bidders" as a result of a $0.00 bid in all eight auctions but were not removed from the sample for analysis because the authors are treating these observations as a choice to "opt out" of the market.

Table 38: WTP Estimates and Associated Price Premiums for Apple Attributes

Sample	Survey Only	Survey and Experiment	
Welfare Measure	Survey Apple WTP	Survey Apple WTP	Auction Apple WTP
WTP LabelS	$1.67/lb	$1.85/lb	$1.49/lb
WTP LabelO	$1.71/lb	$1.89/lb	$1.55/lb
WTP LabelL	$1.79/lb	$1.81/lb	$1.39/lb
WTP LabelT	$1.12/lb	-$1.26/lb	$0.88/lb
Premium(LabelS)	$0.55/lb	$3.12/lb	$0.61/lb
Premium(LabelO)	$0.59/lb	$3.17/lb	$0.67/lb
Premium(LabelL)	$0.67/lb	$3.08/lb	$0.51/lb

Table 39: WTP Estimates and Associated Price Premiums for Steak Attributes

Sample	Survey Only	Survey and Experiment	
Welfare Measure	Survey Steak WTP	Survey Steak WTP	Auction Steak WTP
WTP LabelS	$7.61/lb	$7.27/lb	$6.45/lb
WTP LabelO	$7.58/lb	$7.29/lb	$6.67/lb
WTP LabelL	$7.66/lb	$7.40/lb	$5.87/lb
WTP LabelT	$4.95/lb	$0.77/lb	$4.17/lb
Premium(LabelS)	$2.66/lb	$6.50/lb	$2.28/lb
Premium(LabelO)	$2.63/lb	$6.52/lb	$2.50/lb
Premium(LabelL)	$2.71/lb	$6.63/lb	$1.70/lb

From initial comparison of the table columns, several observations are made. First, compar-

ing the hypothetical WTP in the national survey sample with the experiment participant sample yields a different preference ordering for each population on both products. While the national survey participants preferred local apples over organic apples over sustainable apples, the experiment participants preferred organic apples over sustainable apples over local apples. Additionally, the mean hypothetical WTP estimates on all apple labels were slightly higher for the experiment sample than for the national survey sample. The experiment sample results produced a willingness to accept measure on the typically unlabeled apples, leading to much higher price premiums on the credence attribute labels relative to the typical, unlabeled product.

While the national survey sample preferred local steak over sustainable steak over organic steak, the experiment participants preferred local steak over organic steak over sustainable steak. Additionally, the mean hypothetical WTP estimates on all steak labels were slightly lower for the experiment sample than for the national survey sample. However, because the hypothetical WTP for typical, unlabeled steak in the experiment participant population was so low, the resulting price premiums on the three credence labeled products are much higher for the experiment participants.

Now comparing the hypothetical WTP and the non-hypothetical WTP estimates for the experiment participant sample reveals potential evidence of a degree of hypothetical bias as hypothetical WTP estimates are higher for all credence labeled products than the associated mean auction bids. Additionally, the experiment participants bid a positive amount on average for the typical unlabeled apples, even though the hypothetical results suggested a willingness to accept measure instead. The auction bids induce the same preference ordering as the hypothetical WTP data for apples overall.

The experimental auction bids on steak produce a different preference ordering than the hypothetical data. Hypothetically, the experiment participants preferred local steak over organic steak over sustainable steak. However, according to the non-hypothetical auction bids, the participants preferred organic steak over sustainable steak over local steak. Addi-

tionally, the mean auction bid on the typical, unlabeled steak far exceeded the hypothetical WTP estimate on the same product. Although the experiment participants indicated a general disinterest in the typical unlabeled products in the hypothetical survey scenarios, in the non-hypothetical auction situation they consistently valued the typical, unlabeled steak more than no steak at all. This may be partially due to the sensory experience associated with the auction, that is, seeing the "typical unlabeled" steak (as it was identical in size, texture, and color to all other steaks) persuaded participants that, it too, looked "good enough to eat".

Rounding out this analysis with the final pairwise comparison of the hypothetical WTP from the national survey sample with the non-hypothetical WTP from the experiment participant sample yields interesting results. The hypothetical WTP estimates from the national survey sample are much closer to the mean auction bids received in the experiment than the hypothetical WTP estimates from the experiment participants themselves. This result is encouraging, because it suggests a much smaller magnitude of hypothetical bias if it is present, when considering a sufficiently large sample size (n=500 versus n=40). This result could be primarily context-driven. That is, the experiment participants that followed through and came were recruited primarily from a Health Food specialty grocery store (there was a much higher rate of no-shows for participants recruited from the generic grocery store), they were informed that they would be participating in a food study on sustainability, and they took the survey before participating in the auctions. It is possible that these factors led experiment participants to answer the hypothetical shopping scenario questions more dogmatically, so as to appear consistent with their choice of shopping locale, to promote sustainable or organic agricultural products similar to the ones sold in their grocery store of choice, or because they had not viewed the food products yet.

It is worth noting that the price premiums indicated by the mean auction bids were highest for the organic label for both apples and steak, although we did not make any certification claims on any of the products. We cannot be certain whether or not the participants

automatically inferred that the organic label was backed by USDA certification, as that perception information was not elicited at any point.

Hypothesis Testing of Equal WTP Across Samples

A similar bootstrapping technique was utilized to compare hypothetical and non-hypothetical WTP estimates for the experiment participant sample as outlined in Chapter 3.

Table 40: Comparison of Hypothetical and Non-hypothetical WTP: Apple Survey

	Hypothetical WTP		Non-hypothetical WTP	
Label	Average	95% interval	Average	95% interval
WTP LabelS	$1.85/lb	($1.12/lb, $2.61/lb)	$1.49/lb	($1.29/lb, $1.70/lb)
WTP LabelO	$1.89/lb	($1.23/lb, $2.54/lb)	$1.55/lb	($1.32/lb, $1.78/lb)
WTP LabelL	$1.81/lb	($1.01/lb, $2.56/lb)	$1.39/lb	($1.17/lb, $1.61/lb)
WTP LabelT	-$1.26/lb	(-$2.10/lb, $0.87/lb)	$0.88/lb	($0.73/lb, $1.04/lb)

Table 41: Comparison of Hypothetical and Non-hypothetical WTP: Steak Survey

	Hypothetical WTP		Non-hypothetical WTP	
Label	Average	95% interval	Average	95% interval
WTP LabelS	$7.27/lb	($5.86/lb, $8.70/lb)	$6.45/lb	($5.33/lb, $7.57/lb)
WTP LabelO	$7.29/lb	($5.91/lb, $8.81/lb)	$6.67/lb	($5.52/lb, $7.81/lb)
WTP LabelL	$7.40/lb	($5.98/lb, $8.96/lb)	$5.87/lb	($4.83/lb, $6.92/lb)
WTP LabelT	$0.77/lb	($0.16/lb, $1.37/lb)	$4.17/lb	($3.25/lb, $5.08/lb)

The bootstrapped estimates reveal that the only 95% confidence intervals that do not overlap between hypothetical and non-hypothetical estimates are the WTP measures on typical, unlabeled steaks. This result suggests that the context in which the participants took the survey and the cheap talk script had the effect of mitigating hypothetical bias in the experiment participant sample for the three credence labels. That is, a comparison of 95% confidence intervals leads us to the conclusion that the hypothetical WTP estimates were not significantly different than the non-hypothetical WTP bids from the auction procedure, with the exception of the downward bias on the hypothetical valuations of the typical, unlabeled steak. The evidence of downward bias in the hypothetical typical steak purchases

is in contradiction with accepted theory. Usually, hypothetical bias exhibits a positive trend towards hypothetical responses, that is, $H_0 : WTP_{Hyp} > WTP_{Non-hyp}$.

Tobit Model

The results of the base linear tobit model on non-hypothetical auction bids indicate that all dummy variables corresponding to production label are statistically significant and positive. The tobit MLE coefficients should be interpreted as the magnitude of the change in bid corresponding to a one unit increase in the explanatory variable. Therefore, the coefficient on each of the production labels can be loosely interpreted as the potential price premium captured by that respective labeling scheme. The regression coefficients are consistent with the summary of mean bids, yielding a preference ordering of Organic > Sustainable > Local for both food products. In the apple auction, only age and number of children were statistically significant demographic explanatory variables. Age had a very small, yet positive relationship with bids. Number of children has a slightly larger, negative, relationship with bids. In the steak auction all included demographic variables were statistically significant with the exception of age. Women, generally, bid higher than men, income and education both had moderate positive effects on bids and again the number of children in a household had a negative relationship with auction bids.

Table 42: Tobit Regression Results for Apple Bids

Explanatory Variable	MLE Coefficient	P-value
Gender	-0.12	-0.30
Age	0.01 *	0.047
Income	-0.02	0.542
Education	-0.01	0.830
Children	-0.16**	0.01
Label S	0.67**	0.00
Label O	0.73**	0.00
Label L	0.55**	0.01

Table 43: Tobit Regression Results for Steak Bids

Explanatory Variable	MLE Coefficient	P-value
Gender	-2.98**	0.00
Age	0.04	0.22
Income	0.46*	0.04
Education	0.56*	0.03
Children	-1.18**	0.00
Label S	2.58**	0.00
Label O	2.82**	0.00
Label L	2.07**	0.02

Conclusions

This paper discusses the results from a series of experimental food auctions, used to examine consumer responses to a variety of production labels on apples and ribeye steaks. This work also examines the extent of hypothetical bias in estimates by comparing hypothetically elicited WTP with non-hypothetical auction bids for the same population. Overall, this study consistently found that food labeled as sustainable, organic, or local influenced participant willingness to pay.

It was confirmed using a LR test that the experiment participant sample preferences in the hypothetical shopping scenarios contained in the survey were significantly different than the preferences of the national survey sample. This could be attributed to a number of factors, including difference in demographic make-up and context in which the survey was taken. Additionally, sign reversals on the price coefficient and typical label coefficient were observed for the experiment participant sample on hypothetical preferences.

Pairwise comparisons of the hypothetical WTP from the national survey sample, the hypothetical WTP from the experiment participant sample, and the non-hypothetical WTP bids from the experiment participants yielded interesting results. Across samples the preference ordering on credence labels was different. Specifically, the national survey sample preferred local apples over organic apples over sustainable apples, while the experiment participant sample preferred organic over sustainable over local in both hypothetical and non-hypothetical treatments for apples. The national survey sample preferred local steaks over sustainable steaks over organic steaks, while the experiment participant sample preferred local over organic over sustainable in the hypothetical treatment and organic over sustainable over local in the non-hypothetical treatment for steaks.

A comparison of the hypothetical versus non-hypothetical WTP estimates for the experiment participant sample gave insufficient evidence of positive hypothetical bias, although all hypothetical WTP estimates were higher than the mean WTP bids from the auction on both products. Additionally, the experiment participant sample valued the typical, unla-

beled food products much less in the hypothetical shopping scenarios than in the auctions. The hypothetical WTP from the national survey sample was much closer to the auction bids than the hypothetical WTP from the experiment participants themselves. This result could be indicative of the much smaller sample size and is encouraging evidence that hypothetical bias dissipates slightly for larger, more demographically diverse samples.

The price premiums indicated by mean auction bids were highest for the organic label on both apples and steak, although no certification claims were made on any auctioned products. We cannot be sure if the experiment participants inferred USDA certification from the organic label or not. If the organic label was perceived to be endorsed by the USDA, our results suggest that the USDA Organic certification translates a significantly higher value to consumers. While, sustainable production claims may not yield any more purchase value than organic certification or local production, sustainable labels appear to successfully differentiate food products from their conventional or unlabeled alternatives. This work provides evidence that third-party sustainable certification has potential profitability for producers, especially if USDA Organic certification is prohibitively expensive or otherwise intangible. Furthermore, these results suggest that consumers do not distinguish between locally grown and sustainably grown food products, inferring that the sustainable food label may not be effectively communicating value beyond traditional Òeconomic sustainabilityÓ and the support of local economies. Thus, measuring other environmental and social consequences of sustainable production practices may be worth investigating further for their marketing potential.

Finally, our paper supports the findings in the field literature that sustainable, organic, and local food labels produce positive price premiums on food products. However, due to the makeup of our participant population, we do not suggest taking the price premiums estimated here to be representative of the general United States, or Lansing, MI population. Instead, we hope this study has produced insight into the relationship between these three credence-labeling schemes and relative consumer preferences in the market. Additionally, we

consider this a contribution to the growing literature on experimental auction applications and methods.

APPENDIX

Survey Instrument

This is a survey designed to obtain information from consumers regarding food consumption habits and related issues. Your participation in this survey is entirely voluntary and your responses will be kept in strict confidence. The data gathered from this survey is needed for analysis in the dissertation work of an MSU graduate student. If you have any questions, comments, or concerns regarding this survey, please contact Dr. Robert Shupp by email (shupprob@anr.msu.edu) or by phone (517-432-2754).

Notes for Coders: *Please ensure the following are incorporated into the online survey. Thank you.*

1. Please randomize the order of statements in question number 13.
2. Please randomize the order of the issues in question number 14.
3. Each participant should randomly receive either Block A or Block B of questions numbered 15 through 20. Please track which Block was received by each respondent.
4. Please randomize the order of issues in questions 15 - 20 for each Block.
5. Please randomize the order of the issues in question number 21.
6. Please randomize the value of X in question 24 between 0 and 100. Please replace equation with monetary value once X has been chosen.
7. Please randomize the order of the issues in question number 25.
8. Each participant should randomly receive either Block A or Block B of the questions numbered 26 through 34. Please track which was received by each respondent.
9. Please randomize the order of issues in questions 26-34 for each Block.
10. Moreover, for all questions with randomized designs please create data variables tracking the order of presentation for each respondent.

1. I am:

   ☐ Male

   ☐ Female

2. I am ____ years old. (Please fill in the blank.)

3. I live in the _____ zip code area (Please fill in the blank.) and my annual pre-tax, household income is:

   ☐ Less than $20,000

   ☐ $20,000 - $39,999

   ☐ $40,000 - $59,999

   ☐ $60,000 - $79,999

   ☐ $80,000 - $99,999

   ☐ $100,000 - $119,999

   ☐ $120,000 - $139,999

   ☐ $140,000 - $159,999

   ☐ $160,000 or more

4. The best description of my educational background is:

   - ☐ Did not graduate from high school
   - ☐ Graduated from high school, Did not attend college
   - ☐ Attended College, No Degree earned
   - ☐ Attended College, Associates or Trade Degree earned
   - ☐ Attended College, BachelorâĂŹs (B.S. or B.A.) Degree earned
   - ☐ Graduate or Advanced Degree (M.S., Ph.D., M.D., Law School)
   - ☐ Other (Please explain.) _____

5. There are ___ adults and ___ children living in my household. (Please fill in the two blanks.)

6. What best describes your race?

   - ☐ White or Caucasian
   - ☐ Black or African-American
   - ☐ Asian or Pacific Islander
   - ☐ Mexican or Latino
   - ☐ American Indian
   - ☐ Other (Please describe.)_____

7. Are you the primary food/grocery shopper in your household? (e.g. the individual most often purchasing food) ?

   - ☐ Yes
   - ☐ No

8. Do you currently receive food assistance?

   ☐ Yes

   ☐ No

9. If you receive food assistance, please check off all forms of food assistance you receive:

   ☐ Food Stamps

   ☐ WIC

   ☐ Project Fresh

   ☐ Food Bank

   ☐ Soup Kitchen

   ☐ Other (Please specify.) _____

10. Most households consume on average 21 meals in a typical week. How many of these 21 meals consumed by your household normally include:

    a. Apples: _____

    b. Other fruits, besides apples: _____

11. Consumers purchase food from many sources. Please allocate the proportion of your householdâĂŹs total food purchases across each of the following (Sum to 100%):

    a. Supermarket Retailers (e.g. Kroger, Meijer, Wal-Mart)_____ %

    b. Targeted Retailers (e.g. Whole Foods, Foods for Living, Better Health)_____ %

    c. Food Cooperative (e.g. ELFCO)_____ %

    d. Convenience Stores (e.g. Quality Dairy, 7-Eleven)_____ %

    e. Farmers Markets/ CSAs_____ %

    f. Butcher_____ %

    g. Other (Please describe.) _____

12. Please rate the following statement in terms of your **agreement**:

    "I rarely think about the sustainability of production practices when making food purchases".

Strongly Disagree						Strongly Agree
1	2	3	4	5	6	7

13. Please rate the following statements in terms of your **agreement** (Please circle only one in each row.): "I believe that"

	Strongly Disagree						Strongly Agree
farmers face lower levels of productivity when using sustainable food production practices.	1	2	3	4	5	6	7
farmers could sustainably produce food without much additional monetary expense.	1	2	3	4	5	6	7
farmers currently participate in sound sustainable food production practices.	1	2	3	4	5	6	7
organic products are from farms currently participating in sound sustainable food production practices.	1	2	3	4	5	6	7
locally grown products are from farms currently participating in sound sustainable food production practices.	1	2	3	4	5	6	7
smaller farms are more likely to currently participate in sound sustainable food production practices.	1	2	3	4	5	6	7
corporate farms are more likely to currently participate in sound sustainable food production practices.	1	2	3	4	5	6	7
all products labeled sustainably produced are from farms currently participating in sound sustainable food production practices.	1	2	3	4	5	6	7

14. Please rank the following societal issues in order of importance to you (1 being most important and 8 being the least important):

    a. Human Poverty_____

    b. U.S. Health Care System_____

    c. Food Safety_____

    d. Sustainable Agriculture_____

    e. The Environment_____

    f. Financial Well-Being of U.S. Farmers_____

    g. Consumer Food Prices_____

    h. Animal Well-Being and Welfare_____

In the next section you will be asked to choose which aspects of sustainable apple farming are most and least important to you. Please take the time to read the following definitions as related to sustainable production practices to better help you in your responses.

Ground Cover and Area Management Practices are Employed: Adjacent areas are planted with hedgerows, windbreaks, or other low-maintenance plantings to encourage specific beneficial organisms. Within tree rows, ground cover or mulch are selected and maintained to improve soil microbial activity, organic matter levels and nutrient cycling.

Fertilizer and Nutrient Materials are used minimally: Soil quality, including organic matter content, is established at planting and maintained at an optimum level to minimize commercial fertilizer needs.

Little to No Chemical Pesticides are Used for Pest Management: Chemical pesticides are not used. Alternative strategies are employed, including biopesticides, mating disruption, trap out and/or augmentation with beneficial organisms.

Little to No Chemical Herbicides are Used for Weed Management: Soil quality and ground cover in the orchard and adjoining areas are planned and managed to prevent weeds and weed seed immigration into the orchard. Cultural, mechanical or biological methods are used to control weeds.

Pollinator Management is Employed: Bees are not placed in the orchard until blossoms are open. Pesticides hazardous to bees are not used, or if needed in an emergency, are applied such that they are not hazardous to bees.

Other Pests are controlled using preventative measures, and habitat controls: Habitat is modified around orchards to reduce nesting and perching sites for pest birds.

BLOCK A

15. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Little to No Chemical Pesticides are Used for Pest Management	
	Other Pests are controlled using preventative measures, and habitat controls	
	Farm Size is smaller than average and Corporate Involvement is limited	
	Michigan farmers are financially stable	

16. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Little to No Chemical Herbicides are Used for Weed Management	
	Pollinator Management is Employed	
	Farm Size is small and Corporate Involvement is limited	
	Production, distribution, and sale is done locally	
	Michigan farmers are financially stable	

17. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Fertilizer and Nutrient Materials are used minimally	
	Little to No Chemical Pesticides are Used for Pest Management	
	Little to No Chemical Herbicides are Used for Weed Management	
	Pollinator Management is Employed	
	Other Pests are controlled using preventative measures, and habitat controls	
	Farm Size is small and Corporate Involvement is limited	
	Production, distribution, and sale is done locally	
	Consumer food prices are affordable	
	Michigan farmers are financially stable	

18. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Fertilizer and Nutrient Materials are used minimally	
	Little to No Chemical Pesticides are Used for Pest Management	
	Little to No Chemical Herbicides are Used for Weed Management	
	Other Pests are controlled using preventative measures, and habitat controls	
	Farm Size is small and Corporate Involvement is limited	
	Production, distribution, and sale is done locally	

19. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Fertilizer and Nutrient Materials are used minimally	
	Pollinator Management is Employed	
	Other Pests are controlled using preventative measures, and habitat controls	
	Farm Size is small and Corporate Involvement is limited	
	Consumer food prices are affordable	
	Michigan farmers are financially stable	

20. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Fertilizer and Nutrient Materials are used minimally	
	Farm Size is small and Corporate Involvement is limited	
	Production, distribution, and sale is done locally	
	Consumer food prices are affordable	

BLOCK B

15. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Fertilizer and Nutrient Materials are used minimally	
	Little to No Chemical Herbicides are Used for Weed Management	
	Pollinator Management is Employed	
	Other Pests are controlled using preventative measures, and habitat controls	

16. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Little to No Chemical Herbicides are Used for Weed Management	
	Other Pests are controlled using preventative measures, and habitat controls	
	Production, distribution, and sale is done locally	
	Consumer food prices are affordable	
	Michigan farmers are financially stable	

17. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Little to No Chemical Pesticides are Used for Pest Management	
	Pollinator Management is Employed	
	Other Pests are controlled using preventative measures, and habitat controls	
	Production, distribution, and sale is done locally	
	Consumer food prices are affordable	

18. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Fertilizer and Nutrient Materials are used minimally	
	Little to No Chemical Pesticides are Used for Pest Management	
	Pollinator Management is Employed	
	Production, distribution, and sale are done locally	
	Michigan farmers are financially stable	

19. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Ground Cover and Area Management Practices are Employed	
	Fertilizer and Nutrient Materials are used minimally	
	Little to No Chemical Pesticides are Used for Pest Management	
	Little to No Chemical Herbicides are Used for Weed Management	
	Consumer food prices are affordable	
	Michigan farmers are financially stable	

20. Which one of the following aspects of apple farming do you believe are the most and least important in a sustainable apple production system? Please check only one in each column.

Least Important		Most Important
	Little to No Chemical Pesticides are Used for Pest Management	
	Little to No Chemical Herbicides are Used for Weed Management	
	Pollinator Management is Employed	
	Farm Size is small and Corporate Involvement is limited	
	Consumer food prices are affordable	

21. To the best of your knowledge, for each of the apple production practices indicated, production involves which of the following restrictions? Check all that apply.

	Sustainable	Organic	Local	Typical
Require ground cover and area management				
Limited use of fertilizer and nutrient materials				
Prohibited use of herbicides for crops				
Prohibited use of insecticides for crops				
Prohibited use of genetically modified organisms				
Limited corporate involvement				
Limited farm acreage				

22. Which one of the following best describes the proximity from your home you consider âĂIJlocally grownâĂİ apples to originate from?

    a. within 10 miles

    b. within 50 miles

    c. within 100 miles

    d. within 500 miles

    e. within Michigan

    f. Other (Please describe.) _____

23. Would you be willing to pay a premium for apples labeled as âĂIJsustainably produced?âĂİ

    ☐ Yes

    ☐ No

24. If Yes, Would you buy apples labeled as âĂIJsustainably producedâĂİ if it cost $1.49 + 1.49*X% ?

    ☐ Yes

    ☐ No

24b. If No, Would you buy apples labeled as âĂIJsustainably producedâĂIJ if it cost $1.49 - 1.49*X%?

- ☐ Yes
- ☐ No

25. Please rank the emphasis you believe **farmers** producing âĂIJsustainably produced applesâĂİ place on the following issues/concerns.
(1 being most emphasis and 6 being the least emphasis):

    a. Profitability of their farm

    b. Food prices faced by consumers

    c. Quantity of their farmâĂŹs production

    d. Impact their farm has on the environment

    e. Food safety of products produced on their farm

    f. Quality of apples produced

Please consider the following information [provided by the United States Department of Agriculture]:

The word âĂIJsustain,âĂİ from the Latin sustinere (sus-, from below and tenere, to hold), to keep in existence or maintain, implies long-term support or permanence. As it pertains to agriculture, sustainable describes farming systems that are capable of maintaining their productivity and usefulness to society indefinitely. Such systems must be resource-conserving, socially supportive, commercially competitive, and environmentally sound. Sustainable agriculture was addressed by Congress in the 1990 Farm Bill and the Food, Agriculture, Conservation, and Trade Act of 1990 (FACTA). Under that law, the term sustainable agriculture means an integrated system of plant and animal production practices having a site-specific application that will, over the long term:

1. Satisfy human food and fiber needs
2. Enhance environmental quality and the natural resource base upon which the agricultural economy depends
3. Make the most efficient use of nonrenewable resources and on-farm resources and integrate, where appropriate, natural biological cycles and controls
4. Sustain the economic viability of farm operations
5. Enhance the quality of life for farmers and society as a whole

Consumers are increasingly demanding on the ethical dimension of food quality. This relates to the process of production and trade and its broad impacts on society and the environment. It includes a wide range of social, environmental or cultural issues such as the treatment of workers, a fair return to producers, environmental impacts and animal welfare. Guidelines about what specific practices meet long-term environmental, economic

and social goals and constitute sustainable agriculture is still under debate. However, a handful of groups have attempted to develop standards and/or provide certification services based upon their own guidelines.

In the next section you will be presented with multiple different alternative apples that could be available for purchase in a retail store where you typically shop. Besides the attributes listed below, each product possesses the same characteristics (e.g., similar color and freshness) and is produced in the U.S. Prices vary for each product and are all in $/lb. units. Please consider the following information to help you interpret alternative products.

Label: The display that contains the apples for your purchase may be labeled as follows:

- Sustainable: These apples were produced using sustainable practices.
- Organic: These apples were produced using organic practices.
- Local: These apples were produced for distribution and sale locally.
- Typical: These apples are not labeled to suggest they were produced using any of the criteria listed above.

Certification: The typical product has no certification label. Each labeled product can be certified in one of three ways:

- USDA: The processes used and all claims made by the product label have been verified by the USDA.
- Independent Third Party: The processes used and all claims made by the product label have been verified by a third party unrelated to the farm of origin or retailer.
- Self: The processes used and all claims made by the product label have been verified by the farmer producing the food.

The experience from previous similar surveys is that people often state a higher willingness to pay than what one is actually willing to pay for the good. For instance, a recent study asked people whether they would purchase a new food product similar to the one you are about to be asked about. This purchase was hypothetical (as it will be for you) in that

no one actually had to pay money when they indicated a willingness to purchase. In the study, 80% of people said they would buy the new product, but when a grocery store actually stocked the product, only 43% of people actually bought the new product when they had to pay for it. This difference (43% vs. 80%) is what we refer to as hypothetical bias.

Accordingly, it is important that you make each of your upcoming selections like you would if you were actually facing these exact choices at a store; noting that buying a product means that you would have less money available for other purchases.

BLOCK A

26. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Private 3rd Party	Self	USDA		
Price	\$0.99 /lb	\$0.99 /lb	\$1.99 /lb	\$1.99 /lb	
	☐	☐	☐	☐	☐

27. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Private 3rd Party	USDA	USDA		
Price	\$1.99 /lb	\$1.49 /lb	\$0.99 /lb	\$0.99 /lb	
	☐	☐	☐	☐	☐

28. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	3rd Party	Self	3rd Party		
Price	\$1.49 /lb	\$1.99 /lb	\$0.99 /lb	\$1.49 /lb	
	☐	☐	☐	☐	☐

29. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Private 3rd Party	Private 3rd Party	Self		
Price	$1.49 /lb	$1.49 /lb	$1.99 /lb	$0.99 /lb	
	☐	☐	☐	☐	☐

30. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Self	Self	Self		
Price	$0.99 /lb	$1.49 /lb	$0.99 /lb	$1.49	
	☐	☐	☐	☐	☐

31. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	USDA	Self	Private 3rd Party		
Price	$1.99 /lb	$1.49 /lb	$1.99 /lb	$1.99 /lb	
	☐	☐	☐	☐	☐

32. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	USDA	Private 3rd Party	Self		
Price	$1.99 /lb	$1.99 /lb	$0.99 /lb	$1.99 /lb	
	☐	☐	☐	☐	☐

33. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	USDA	Self	Self		
Price	$1.49 /lb	$0.99 /lb	$1.49 /lb	$0.99 /lb	
	☐	☐	☐	☐	☐

34. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	USDA	USDA	Private 3rd Party		
Price	$0.99 /lb	$0.99 /lb	$0.99 /lb	$0.99 /lb	
	☐	☐	☐	☐	☐

BLOCK B

26. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Self	Self	USDA		
Price	$1.99 /lb	$1.99 /lb	$1.49 /lb	$0.99 /lb	
	☐	☐	☐	☐	☐

27. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Private 3rd Party	USDA	Self		
Price	$0.99 /lb	$1.99 /lb	$1.49 /lb	$1.99 /lb	
	☐	☐	☐	☐	☐

28. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	3rd Party	3rd Party	3rd Party		
Price	$1.99 /lb	$0.99 /lb	$1.49 /lb	$1.49 /lb	
	☐	☐	☐	☐	☐

29. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	USDA	Private 3rd Party	USDA		
Price	$0.99 /lb	$1.49 /lb	$1.49 /lb	$1.49 /lb	
	☐	☐	☐	☐	☐

30. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	USDA	USDA	USDA		
Price	$1.49 /lb	$1.99 /lb	$1.99 /lb	$1.49 /lb	
	☐	☐	☐	☐	☐

31. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Self	Private 3rd Party	USDA		
Price	$1.49 /lb	$0.99 /lb	$0.99 /lb	$1.99 /lb	
	☐	☐	☐	☐	☐

32. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Self	USDA	Self		
Price	$1.99 /lb	$0.99 /lb	$1.99 /lb	$1.49 /lb	
	☐	☐	☐	☐	☐

33. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Self	3rd Party	3rd Party		
Price	$0.99 /lb	$1.99 /lb	$1.99 /lb	$0.99 /lb	
	☐	☐	☐	☐	☐

34. Which one of the following apple displays listed below would you choose to purchase from?

Label	Sustainable	Organic	Local	Typical	I choose not to purchase any of these options.
Certification	Self	USDA	Private 3rd Party		
Price	$1.49 /lb	$1.49 /lb	$1.49 /lb	$1.99 /lb	
	☐	☐	☐	☐	☐

Thank you for your time in completing this survey. Your input will strengthen our research and help us obtain more accurate conclusions. If you wish to add any comments please feel free to do so here:

REFERENCES

REFERENCES

Alfnes, Frode, and Kyrre Rickertsen. "European Consumers' Willingness to Pay for U.S. Beef in Experimental Auction Markets." American Journal of Agricultural Economics 85.2 (2003): 396-405.

Briggeman, Brian C., and Jayson L. Lusk. "Preferences for Fairness and Equity in the Food System." European Review of Agricultural Economics 38.1 (2011): 1-29.

Cameron, A. C. and P.K. Trivedi. Microeconometrics Using Stata. College Station: Stata Press, 2010.

Chern, Wen S., and Chun-Yu Chang. "Benefit Evaluation of the Country of Origin Labeling in Taiwan: Results from an Auction Experiment." Food Policy 37.5 (2012): 511-9. PAIS International.

Clonan, A., M. Holdworth, J. Swift, and P. Wilson, 2010. UK Consumers Priorities for Sustainable Food Purchases. Conference Proceedings. http://purl.umn.edu/91948

Colson, G. and W. E. Huffman. ÒConsumersÕ Willingness to Pay for Genetically Modified Foods with Product-Enhancing Nutritional Attributes.Ó (2011) Amer. J. Agr. Econ. 93(2): 358Ð363.

Corrigan, J. R., and M. C. Rousu. 2011. ÒAre Experimental Auctions Demand Revealing when Values are Affiliated?Ó American Journal of Agricultural Economics 93, January.

Evans, Jason R., et al. "Determining Consumer Perceptions of and Willingness to Pay for Appalachian Grass-Fed Beef: An Experimental Economics Approach." Agricultural and Resource Economics Review 40.2 (2011): 233-50.

Hellyer, Nicole Elizabeth; Fraser, Iain; Haddock-Fraser, Janet. ÒFood Choice, Health Information and Functional Ingredients: An Experimental Auction Employing BreadÓ Food Policy 37. 3 (Jun 2012): 232-245.

Hobbs, Jill E., Kim Sanderson, and Morteza Haghiri. "Evaluating Willingness-to-Pay for Bison Attributes: An Experimental Auction Approach." Canadian Journal of Agricultural Economics 54.2 (2006): 269-87.

Lee, Ji Yong, et al. "Valuing Traceability of Imported Beef in Korea: An Experimental Auction Approach." Australian Journal of Agricultural and Resource Economics 55.3 (2011): 360-73.

Lusk, J. et.al. ÒEffect of Information About Benefits of Biotechnology on Consumer Acceptance of Genetically Modified Food: Evidence from Experimental Auctions in the United

Stated, England, and France.Ó European Review of Agricultural Economics Vol 31 (2) (2004) pp. 179-204.

Lusk, J. et.al. ÒConsumer Welfare Effects of Introducing and Labeling Genetically Modified Food.Ó Economics Letters 88 (2005) 382Ð388.

Lusk, J. and J. F. Shogren, 2007. Experimental Auctions: Methods and Applications in Economic and Marketing Research. Cambridge University Press.

Nayga, R. M., R. Woodward and W. Aiew. ÒWillingness to Pay for Reduced Risk of Foodborne Illness: A Nonhypothetical Field Experiment.Ó Canadian Journal of Agricultural Economics 54 (2006) 461Ð475.

Onken, Kathryn A., John C. Bernard, and Pesek,John D.„Jr. "Comparing Willingness to Pay for Organic, Natural, Locally Grown, and State Marketing Program Promoted Foods in the Mid-Atlantic Region." Agricultural and Resource Economics Review 40.1 (2011): 33-47.

Onozaka, Yuko, Gretchen Nurse, and Dawn Thilmany McFadden. "Defining Sustainable Food Market Segments: Do Motivations and Values Vary by Shopping Locale?" American Journal of Agricultural Economics 93.2 (2011): 583-9.

Onozaka, Yuko, and Dawn Thilmany McFadden. "Does Local Labeling Complement Or Compete with Other Sustainable Labels? A Conjoint Analysis of Direct and Joint Values for Fresh Produce Claims." American Journal of Agricultural Economics 93.3 (2011): 693-706.

Rousu, Matthew C., and Jay R. Corrigan. "Estimating the Welfare Loss to Consumers when Food Labels do Not Adequately Inform: An Application to Fair Trade Certification." Journal of Agricultural and Food Industrial Organization 6.1 (2008)

Saunders, C. et.al., 2010. Consumer Attitudes Towards Sustainability Attributes on Food Labels. Conference Proceedings. http://ageconsearch.umn.edu/handle/96944

Umberger, Wendy J. "U.S. Consumer Preference and Willingness-to-Pay for Domestic Corn-Fed Beef Versus International Grass-Fed Beef Measured through an Experimental Auction." Agribusiness 18.4 (2002): 491-504.

Chapter 5: Summary

Surprisingly little economic research has focused on sustainability in the context of agricultural production practices before this work. Generally, the literature on other value-added food attributes, such as organic, local, natural, or GMO-free have found that credence attributes have a positive impact on consumer valuations and lead to higher WTP, despite the inability to directly identify the attribute either before purchase or after consumption. "Sustainably produced" is an attribute with no absolute definition, leaving sustainable food labels much more open to consumer perceptions and inferences that go beyond the information provided on packaging or marketing labels. Understanding the perceptions and preferences consumers hold regarding what a "sustainably produced" labeling scheme implies is crucial to determining if the product will succeed in the market.

Best-worst analysis was applied in this research to investigate the degree of importance consumers give to ten sustainable farming production attributes and in particular was used to determine behavioral differences across clustered subgroups of the population sample. While consumers are generally familiar with organic standards outlining land-use and environmental impact variables of production, this study indicates that environmental indicators of sustainability are less important to consumers than economic dimensions. Based on our initial results, size, scale, and geographic scope capture the attributes of sustainability that are most important to consumers. Therefore, differentiating food claims on the level of locality provides a marketing avenue worth exploring further, as supported by the growing literature on the local foods movement.

Our study also provides supporting evidence that the term "sustainable" could be causing more confusion than it is adding value in credence labeling schemes. The latent class assessments on both beef and apple consumer samples indicated that 50 - 55% of the evaluated population is highly price-driven or significantly indifferent to varied sustainability attributes. However, these results could also imply that consumers perceive sustainability as a bundle of attributes, and thus ranked them equally important in this setting. Given the

associated heterogeneous economic welfare impacts that would come as a result of across-the-board market adjustments, such as government endorsed sustainability labeling, or bans on select farming practices, this is an important result.

Our analysis of hypothetical purchase choices demonstrates that positive price premiums can be captured by sustainably produced labeling claims, relative to similar unlabeled and conventionally produced food products. However, the price premiums calculated in our model reveal that there is a comparable tradeoff in quality associated between local, organic and sustainably labeled food products. Furthermore, detailed information about sustainable certification guidelines had no significant impact on choosing the sustainably labeled products. Thus, we conclude that, based on our results, consumer demand for sustainably produced food may not be distinctly differentiable from its local and organic counterparts. Overall, these findings suggest that profitable marketing opportunities may exist for firms interested in selling sustainably produced food products, however there needs to be considerable effort put into leveling information asymmetries about product quality if sustainable label claims are to be differentiated from the local food movement or more recognizable organic principles.

The experimental auction data provided consistent evidence that food labeled as sustainable, organic, or local influenced participant willingness to pay. However, it was confirmed using a LR test that the experiment participant sample preferences in the hypothetical shopping scenarios contained in the survey were significantly different than the preferences of the national survey sample. This could be attributed to a number of factors, including difference in demographic make-up and context in which the survey was taken.

A comparison of the hypothetical versus non-hypothetical WTP estimates for the experiment participant sample, using bootstrapping methods, gave insufficient evidence of positive hypothetical bias, although all hypothetical WTP estimates were higher than the mean WTP bids from the auction on both products. Additionally, the experiment participant sample valued the typical, unlabeled food products much less in the hypothetical shopping scenarios

than in the auctions. The hypothetical WTP from the national survey sample was much closer to the auction bids than the hypothetical WTP from the experiment participants themselves. This result could be indicative of the much smaller sample size and is encouraging evidence that hypothetical bias dissipates slightly for larger, more demographically diverse samples.

While, sustainable production claims may not yield any more purchase value than organic certification or local production, sustainable labels appear to successfully differentiate food products from their conventional or unlabeled alternatives. This work provides evidence that third-party sustainable certification has potential profitability for producers, especially if USDA Organic certification is prohibitively expensive or otherwise intangible. Furthermore, these results suggest that consumers do not distinguish between locally grown and sustainably grown food products, inferring that the sustainable food label may not be effectively communicating value beyond traditional Òeconomic sustainabilityÓ and the support of local economies. Thus, measuring other environmental and social consequences of sustainable production practices may be worth investigating further for their marketing potential.

It is hard to conclude whether sustainable food labels will or will not follow the same trajectory as organic. However, it is encouraging to see more emphasis on the differences between organic and sustainable in the food system rhetoric, at both academic and popular levels. Inevitably, more information is becoming available for consumers to distinguish between food labels. Therefore, a clear extension of this work is needed in comparing consumer preferences for sustainably labeled foods with other bundled labels, such as Organic + Local. Furthermore, a study should be conducted on the effect of using "sustainable" attributes as a marketing tool, in mediums other than labels, on consumer valuation.