THREE ESSAYS: VALUATION FOR AND MARKETING SPILLOVER OF DISCOLORED BEEF AND EMPLOYER WILLINGNESS TO PAY FOR RECENT COLLEGE GRADUATE ATTRIBUTES

By

RYAN FEUZ

Bachelor of Science in Economics and Finance
Utah State University
Logan, Utah
2014

Master of Science in Applied Economics
Utah State University
Logan, Utah
2016

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2019

THREE ESSAYS: VALUATION FOR AND MARKETING SPILLOVER OF DISCOLORED BEEF AND EMPLOYER WILLINGNESS TO PAY FOR RECENT COLLEGE GRADUATE ATTRIBUTES

Dissertation Approved:

F. Bailey Norwood

Dissertation Adviser

Derrel S. Peel

John Michael Riley

Ranjith Ramanathan

ACKNOWLEDGEMENTS

Thank you first and foremost to my wife for her enduring support, as well as to my parents. Additionally, thank you to my committee members and advisor Dr. Norwood for their help and support throughout my time at Oklahoma State.

Name: RYAN FEUZ

Date of Degree: JULY, 2019

Title of Study: THREE ESSAYS: VALUATION FOR AND MARKETING SPILLOVER OF DISCOLORED BEEF AND EMPLOYER WILLINGNESS TO PAY FOR RECENT COLLEGE GRADUATE ATTRIBUTES

Major Field: AGRICULTURAL ECONOMICS

Abstract: This dissertation is comprised of two essays focusing on discoloration in retail beef, with the third essay evaluating the value to employers of recent college graduate attributes.

Consumers prefer beef to be bright cherry-red. As beef becomes discolored we expect consumers are willing to pay less for it. The first study's objective is to determine the willingness-to-pay (WTP) for varying levels of discoloration within both beef steak and ground beef relative to non-discolored beef. Two types of discoloration are considered: coverage (percent of surface area discolored) and intensity (the darkness of the discoloration). The estimated WTP discounts can be useful to retailers as they attempt to minimize losses associated with discolored beef.

The second study continues with the theme of discolored beef. Whereas the first study was concerned with the value of the discolored beef, the second study focuses on the value of and preferences for non-discolored beef when marketed together with discolored beef. It is plausible that marketing discolored beef alongside non-discolored beef may have a spillover effect on the value of the non-discolored beef. This study is concerned with determining if a spillover effect exists and determining the direction of the effect.

The third study is unrelated to the first two essays. In this essay we introduce a new method of stated preference valuation called design valuation. This method operates by defining a general good as a collection of attributes and assigning prices to those attributes. Respondents are then asked to design their optimal good based on selecting any combination of those attributes. By varying the attribute prices within and across surveys, the value of each attribute can be inferred. We apply the design valuation method to accomplish the purpose of this study which is to estimate the value employers place on various college graduate attributes. We add to the literature on recent college graduate attribute valuation by further classifying attribute values by employer type. Four employer type classifications are used; those that prefer to hire graduates from an agricultural, business, engineering, or other college.

TABLE OF CONTENTS

Chapter Page

I. DO CONSUMERS HAVE AN APPETITE FOR DISCOLORED BEEF? 1

Abstract 1
Introduction 2
Materials and Methods 4
Beef Items Used in the Choice Experiment 4
Survey Design and Participants 5
The Choice Experiment 8
Hypothetical Bias Correction 10
Attention Checks 11
Statistical Model 14
Results 19
Conditional Logit Estimation 19
Willingness to Pay (WTP) 22
Discussion 25
Failing the Trap Model 25
Strategic Discounts for Quick Sale of Discolored Beef 26
Limitations 30
Conclusions 31

II. THE SPILLOVER EFFECT OF MARKETING DISCOLORED BEEF ON CONSUMER PREFERENCES FOR NON-DISCOLORED BEEF 40

Abstract 40
Introduction 41
Literature on Spillover Effects 43
Methods and Data 45
Study 1 45
Study 2 46
Results 54
Study 1 54
Study 2 56
WTP Estimates 59
Summary and Conclusions 60
Limitations 60
Conclusions 60

Chapter Page

III. INTRODUCING DESIGN VALUATION WITH APPLICATION TO VALUEING COLLEGE GRADUATE ATTRIBUTES 71

Abstract 71
Introduction 72
Theory 76
Data 78
Procedure 84
Results 85
Summary and Conclusions 91

REFERENCES 97

APPENDICES 102

LIST OF TABLES

Table Page

1.1 Summary Statistics of Survey Respondents 32

1.2 Steak and Ground Beef Model Results.....33

1.3 Willingness-to-Pay Estimates for Discoloration Attributes in Steak.....34

1.4 Willingness-to-Pay Estimates for Discoloration Attributes in Ground Beef.....35

1.5 Comparing Demographics of Respondents Who Passes vs. Failed the Trap 36

2.1 Questionnaire Respondent Descriptive Statistics 62

2.2 Color Characteristics of Bright-red and Discolored Steak on Day of Study 63

2.3 Summary of Responses to Questionnaire Statements.....64

2.4 Summary of Factor Analysis Loadings.....65

2.5 Regression Results for Statements 1-4 66

2.6 Results of Combined Statements 1-4 Two Class Finite Mixture Model 67

2.7 Ordered Logit Results for Statements 5 and 6 68

3.1 Employer Demographics (Sample Size = 453).....94

3.2 Value for Recent College Graduate Attributes with Respect to the Type of College the Employer Prefers to Hire From.....95

LIST OF FIGURES

Figure Page

1.1 Overview of Survey 37

1.2 Twelve Levels of Discoloration Used in the Choice Experiment.....38

1.3 Cheap Talk Script/Trap Question 39

2.1 Grid Style Question from Nationwide Consumer Survey 69

2.2 Summary Results of Survey Question Asking About Appeal of Steak.....70

3.1 Design Valuation Question.....96

CHAPTER I

DO CONSUMERS HAVE AN APPETITE FOR DISCOLORED BEEF?

Abstract

Consumers prefer beef to be bright, cherry-red in color. As beef becomes discolored consumers may perceive it unwholesome, reducing their willingness-to-pay (WTP). This study uses a choice experiment in an online survey to measure how discoloring affects consumer WTP for beef, using a cheap talk script combined with a trap question to mitigate hypothetical bias. Results indicate even slightly discolored beef may require at least a 50 percent discount before the average consumer will purchase it. However, if the quantity of discolored meat to be sold is low, retailers can sell discolored meat at lower discounts to less discriminating consumers.

Key Words: Attention bias; Beef color; Beef discoloration; Trap question; Willingness-to-pay

Introduction

Consumer purchasing decisions for beef are influenced by attributes of the meat itself. Among these attributes, the color can greatly influence consumer decisions. Quality inferences are based in part on the color of the meat. Research has shown the preferred color of beef to be bright, cherry-red (Carpenter, Cornforth, and Whittier, 2001; Killinger et al., 2004). Deviations from the preferred color create some level of unacceptability among consumers. As the surface of the meat becomes discolored, turning brownish-red to brown, consumers may perceive this as a condition of unwholesomeness (Faustman and Cassens, 1989). This is expected to result in a decreased willingness-to-pay (WTP) for discolored beef (Grebitus, Jensen, Roosen, and Sebranek, 2013) and, therefore, discolored beef cuts must be marked down in price, faced and repackaged, ground, cooked and sold as prepared foods, or discarded (Smith, Morgan and Tatum, 1993). This creates waste and represents a significant cost to the beef industry (Smith, et al., 2000). Hermel (1993) and Williams, et al. (1992) both estimated that approximately $1 billion is lost to the beef industry annually due to discoloration. To help mitigate this cost, research has focused on ways to prolong the bright cherry-red color through various packaging and display techniques (Greene, Hsin and Zipser, 1971).

There is an abundance of consumer preference research in retail beef. Many attributes have been considered such as tenderness, taste, fat content, and region of production (Lusk et al, 2001; Feuz et al, 2004; Loureiro and Umberger, 2003). Studies measuring WTP that have focused on color, generally have considered only the color of non-discolored beef items (color vibrancy), finding that color does impact purchasing decisions, with bright, cherry-red beef being preferred (Carpenter, Cornforth, and Whittier, 2001; Killinger, et al, 2004).

Grebitus, Jensen, and Roosen (2013) evaluated consumer acceptance of modified atmosphere packaging (MAP) as well as consumer WTP for color attributes in ground beef. The color attributes they evaluated varied from bright cherry-red, light red, and brownish red. Similar to other past research, they found the preferred color to be bright, cherry-red and estimated the average WTP for cherry-red ground beef was $2.00/lb. more than for brownish red ground beef. Within their choice experiment the brownish-red color attribute was constant and did not have varying levels. Additionally, the researchers did not provide any description of the level of discoloration beyond stating that the color was a consistent brownish-red achieved by aerobically packaging the ground beef in permeable overwrap and irradiating it with 1 kGy to achieve a standardized and consistent metmyoglobin-like brownish red color to represent a meat color that has begun to deteriorate in retail display. The extent to which consumers will purchase varying levels of discolored beef at discount prices has not been fully evaluated.

The objective of this study is to determine the WTP for varying levels of discoloration within both beef steak and ground beef relative to non-discolored beef. Using an internet survey, subjects are shown pictures of different beef items that are identical save for their level of discoloration. The pictures are embedded in a series of choice experiment choice sets, similar to those in Gracia and de-Magistris (2013) where each beef product is given a price and subjects are asked which item they would most likely purchase in a grocery store. By evaluating their choices under different price scenarios it is possible to identity the discount one would have to apply to a discolored beef item before the average consumer would purchase it instead of a similar non-discolored item.

This paper is organized as follows. The next section describes the beef items, participants, and survey instrument used to measure WTP for discolored beef. Also included in the second section is a discussion of the statistical model used and the survey techniques employed to correct for hypothetical bias and inattentiveness. The third section presents the results, and the fourth section provides a discussion of what the results imply about consumers' true WTP for discolored beef and the discounts retailers would need to assign them for sale. The fifth section concludes.

Material and Methods:

Beef items used in the choice experiment

Two beef items, steak and ground beef, were evaluated. The items were purchased fresh and then intentionally discolored so that pictures could be taken of the same beef item at different levels of discoloration, allowing us to evaluate consumer preferences for beef holding all attributes constant except the level of discoloration and price.

The steak came from market-age cattle purchased from a local purveyor. The approximate postmortem age was 14 days. A 2.5-cm-thick steak was cut from the ribeye steak section using a meat slicer (Bizerba USA Inc., Piscataway, NJ) and placed onto foam trays with absorbent pads, over-wrapped with a PVC film (oxygen-permeable polyvinyl chloride fresh meat film; 15,500–16,275 cm3 O2/m2/24 h at 23 °C, E-Z Wrap Crystal Clear Polyvinyl Chloride Wrapping Film, Koch Supplies, Kansas City, MO). After packaging, the steak was placed in a coffin-style open display case maintained at 2°C ± 1 under continuous lighting for eight days (1612 to 1800 lx, Philips Delux Warm White Fluorescent lamps; Andover, MA; color rendering index = 86; color temperature = 3000 K). During these eight

days the steak became increasingly discolored, and was documented using a digital Cannon SLR camera. The 10 percent fat ground beef item was managed in an identical fashion except it required only five days of exposure. The discoloration gradually increased, as it naturally would in a retail environment. Photo editing software was then used to vary the intensity (level of darkness) of the discoloration.

Survey Design and Participants

The survey was created within the online survey creation platform, Qualtrics (Qualtrics 2017) and administered by Survey Sampling International (Survey Sampling International, LLC 2017). It was administered September 6, 2017 through October 3, 2017. In total 2,598 respondents completed the survey with the target sample set to be representative of the U.S. consumer population. Summary statistics for the demographics of the survey sample can be seen in Table 1.1

By comparing data from the United State Census Bureau (2018) it can be seen that overall the sample 'gender' and 'income' demographics are similar to the U.S. population. In 2017 the U.S. was estimated as 50.8 percent female with a median income of $55,322. Age and education attainment in the sample may be slightly misrepresented. In 2017, 87 percent of the U.S. population (age≥25) had an education level of high school graduate or higher while 30.3 percent had a Bachelor's degree or higher. This would indicate that our sample has a slightly higher average education level than the estimated U.S. population. The median age of the US population is 37.7 which is consistent with our sample. However, our sample appears to have a slightly greater percentage of those in the 18-25 age group as well as the 50-65 group.

The survey consisted of five main sections, shown in Figure 1.1. The first section confirmed that the respondents' computer monitor display (the survey was not allowed on smartphones or tablets) and the subjects' eyesight allowed them to distinguish between greater and lesser amounts of intensity (level of darkness) of discoloration. This was accomplished by presenting them with three steaks with different intensity levels and asking them to select the steak with the darkest discoloration, and then repeating the question for ground beef, only asking them to select the ground beef that has the lightest discoloration. In total 71.7 percent of respondents correctly identified the steak with the darkest intensity of discoloration while only 50.8 percent correctly identified the ground beef that had the lightest intensity. Overall, this demonstrates that respondents were often able to distinguish differences in the level of intensity of discoloration.

It was somewhat unexpected that a smaller percentage of respondents correctly answered the question with regards to ground beef as compared to steak. The steak question was asked first so while it is possible that it was more difficult for respondents to notice a difference in intensity within ground beef, it could also be the case that some respondents didn't notice the change in question format from 'darkest' to 'lightest' discoloration. It would then be plausible that some level of inattentiveness could explain the lower percentage of respondents identifying the correct ground beef. While we considered dropping respondents who failed to correctly identify the darkest or lightest discolored beef item, we have chosen to retain them within the sample. Failing to identify the difference in intensity once does not guarantee they could not have made the correct selection in subsequent attempts. Additionally, as this survey was administered online and viewed through individual participant computer monitor displays there is no way of knowing whether the individuals

answered incorrectly because of inattentiveness or poor eyesight, or if their monitors failed to accurately portray the subtle differences in intensity of discoloration. As some shoppers may have poor eyesight in the store as they might in the survey, to best represent the population of beef consumers the individuals who did not distinguish between intensity levels of discolored beef in the preliminary questions were retained in the sample. The decision to retain these individuals was made before any statistical estimations were performed.

The second section of the survey contained the first set of choice experiment questions. There was one set of questions for steaks and another section for ground beef. This section randomly choose whether to display thirteen choice experiment questions for the steaks or the ground beef. Each choice set question contained three beef products with varying discoloration, sold at prices which were (overall) orthogonal to the amount of discoloration. At each choice set question the subject was asked to indicate which of the three items they would mostly likely purchase in a store. This was a hypothetical choice, and studies have shown that in such cases subjects usually overestimate the actual amount they are willing to pay (Penn and Hu, 2018). Fortunately, there are tools available for reducing this hypothetical bias.

The third section then presented a version of cheap talk script popular among researchers for reducing hypothetical bias (Penn and Hu, 2018). Of course, a cheap talk script is only effective if the respondent reads it, and it is well known that respondents sometimes skim information or even answer questions without reading them well. Thus, a trap question was embedded within the cheap talk script to confirm whether the subject read the script. To test whether the cheap talk script/trap question has an effect it was only seen by half of the respondents.

The fourth section of the survey provided another thirteen choice set questions from the choice experiment using a different beef product than in the second section. For example, Figure 1.1. shows a case where thirteen choice set questions for steaks were provided in the second section, in which case the fourth section contained an equal number of choice set questions for ground beef. The last section asked a number of questions about the subjects' beliefs about ground beef and steak and collected demographic data.

The Choice Experiment

The second and fourth section of the survey asked respondents to respond to a number of choice set questions in the choice experiment. The choice experiment method involves asking respondents to state their preferences over hypothetical alternative scenarios, goods or services. Each alternative in the choice sets is then described by several attributes which can vary at multiple levels and the participants' responses are used to determine whether preferences are significantly influenced by the attributes. Each choice set in our survey required respondents to evaluate three beef products sold at different prices and indicate which one they would most likely purchase. Aside from price, the beef products varied by level of discoloration. Two types of discoloration were used: coverage (percent of surface area discolored) and intensity (the darkness of the discoloration). It is possible that one type of discoloration elicits a greater negative reaction, and that there are interaction effects between the two types. For this reason the beef products were allowed to differ in both their coverage and intensity of discoloration.

Figure 1.2. illustrates the twelve different levels of discoloration used in the choice experiment, using the steaks as an example (a similar figure could be made for ground beef).

The coverage attribute could take the values: 25, 50, 75, or 100 percent. The intensity of the discoloration was either light, medium, or dark. Each beef product in a choice set could be thus described in terms of its coverage, intensity, and price. Twelve indicator variables were used to represent each possible combination of coverage and intensity of discoloration possible.

A price was also assigned to each beef product in a choice set. For steaks the price (per pound) could take one of the five values: $5.25, $7.25, $8.25, $9.25, or $11.25. For ground beef the price (per pound) could take the values $2.00, $3.00, $3.50, $4.25, or $6.25. An experimental design procedure was employed to construct choice sets such that the price assigned to each ground beef product was independent of its discoloration coverage or intensity, and to ensure efficient parameter estimates of the discoloration variables.

With both the steak and ground beef portions of the survey having one attribute varying at five levels (price), one attribute varying at four levels (discoloration coverage), and one attribute varying at three levels (discoloration intensity) the total number of unique steak/ground beef descriptions was 60 (51 x 41 x 31). A full factorial design would have required a lengthy choice set and would not have been feasible for this survey. To shorten the choice set required and still ensure an efficient design, the choice sets in both the steak and ground beef portions of the survey were created using a fractional factorial design. This was accomplished through use of the Proc Optex procedure within SAS software. This procedure allows for much smaller choice sets to be used while still ensuring an efficient design is generated by maximizing the D-efficiency score. The resulting design required 13 questions for the steak and 13 questions for the ground beef portion of the survey. To help safeguard against the possibility of earlier questions affecting consumers' choices to later questions in a

systematic way, the order of choice sets was randomized, as was the question order within choice sets.

Hypothetical Bias Correction

This choice experiment relied on stated preference methodology. Concerns are often raised with the use of stated preference methods to assess consumer WTP for nonmarket goods. Perhaps the largest concern comes from recognizing a lack of consistency that can exist between hypothetical behavior and behavior with real economic consequences (Penn and Hu, 2018). There are several methods both ex ante and ex post for correcting for this potential hypothetical bias. Examples of ex post methods of hypothetical bias correction often include calibration techniques (Blackburn, Harrison, and Rutström, 1994; Fox et al., 1998), as well as uncertainty coding (Champ et al., 1997).

One ex ante method known as 'cheap talk script' attempts to address hypothetical bias by directly informing respondents of the nature of the problem and urging them to actively consider the hypothetical nature of the questions and to respond with what they would choose in a real market situation characterized by actual payment of money. The cheap talk script method was initially implemented by Cummings and Taylor (1999) and has become increasingly popular. Loomis (2014) provides a thorough summary of various ex ante and ex post methods of hypothetical bias correction. In reviewing eight studies that used cheap talk script Loomis (2014) finds three of the eight effectively eliminate the bias, three reduce the bias, one has no effect, and one study appears to over-correct for the hypothetical bias.

As Loomis (2014) notes, there is still no singular solution agreed upon to correct for hypothetical bias. What is agreed upon is that it can be significant and some attempt should be made to control for it. As cheap talk scripts help reduce hypothetical bias in most studies, and because they are easy to include in a survey, this study employed a cheap talk script. However, as shown in Figure 1.1. it did so for only one-fourth of the steak and one-fourth of the ground beef choice set questions, allowing us to test whether it reduces WTP in the survey.

Attention Checks

As with nearly all studies involving survey methodologies, measurement error arising from the lack of respondent attentiveness is also a concern. Respondent attentiveness has long been a concern in survey research. Even 50 years ago Cannel and Kahn (1968) maintained that when the optimal length for a survey is surpassed respondents become less motivated to respond, put forth less cognitive effort, and may skip questions altogether, leading to a decrease in the quality of the data gathered. Krosnick (1991) describes the tendency for survey respondents to lose interest and become distracted or impatient as they progress through a survey, putting less and less effort into answering questions. Krosnick refers to this tendency as 'satisficing'. Following Krosnick, many researchers began developing ways to recognize respondents who had potentially engaged in satisficing and, therefore, were inattentive in their responses.

In general, as respondents become less engaged, the tendency is to begin satisficing rather than attentively answering questions. This can result in more acquiescent responding, more frequent selection of non-substantive responses such as 'don't know', non-

differentiation in rating scales, choosing the first listed response (i.e. primacy), endorsing the status quo instead of endorsing social change, failing to differentiate among a set of diverse objects in ratings, and random responding. Such behaviors degrade the quality of data and diminish the ability of researchers to base strong conclusions on resulting studies. The online platform for survey administration is far from immune from these problems (Heerwegh and Loosveldt, 2008; Malhotra, 2008).

There are several methods utilized within the literature to try to identify inattentive respondents. Examples of such methods include trap questions, reverse worded questions, identifying respondents who give straight-lined responses to gridded matrix questions, excessive selection of "I don't know," nonsensical responses to open-ended questions, and short response times. This study focuses on the use of a trap question as a method of inattentive respondent control.

Oppenheimer, Meyvis, and Davidenko (2009) explore the use of a trap style question specifically known as an instructional manipulation check (IMC) as a way to identify inattentive respondents, remove the noise they cause in the data, and increase reliability of the dataset while simultaneously increasing the statistical power of results. A typical IMC consists of a question embedded within the survey that may appear to be similar to other questions in length and response format, however, unlike the other questions, the IMC asks participants to ignore the standard response format and instead provide a confirmation they have read the instructions. If respondents have carefully read the question and followed the instructions correctly then this indicates they have been attentive in the survey process. In contrast, if respondents do not follow instructions correctly for the IMC this indicates they did not read the question carefully and provides an indirect measure of satisficing.

Oppenheimer, Meyvis, and Davidenko (2009) are able to demonstrate through the use of an IMC that they successfully identify satisficing participants and by eliminating those respondents' responses from the dataset are able to increase the statistical power and reliability of a dataset.

Once inattentive respondents have been identified, often the common practices is to eliminate their responses from the sample. However, Berinsky, Margolis, and Sances (2014) caution against this practice. They assert if attentive and inattentive respondents are different types of people, removing all inattentive respondents may bias the sample. If attention on a survey is a function of the characteristics of respondents, then dropping respondents who fail the trap question can skew the demographics of the sample. Using OLS, Berinsky, Margolis, and Sances (2014) regress many characteristics of the respondents on the binary dependent variable of whether they passed the trap, and show those who pass do share common characteristics and are thus not randomly distributed throughout the survey sample demographics. They conclude their paper by offering three suggestions in the use of screener (trap) questions. First, the use of a single screener question is often insufficient. Passing the screener at one point in time does not imply passage at another point in time by the same respondent. Second, because screener passage is in part a function of measurable demographic characteristics, researchers should not simply discard respondents who fail screeners. Instead, results should be stratified by attention. Screeners have been shown to be successful in reducing noise in a sample. However, discarding all who fail can potential skew sample demographics. They recommend transparency when presenting results as a way to reconcile these two points. Third, researchers should analyze predictors of screener passage

in their sample. This allows for researchers to have a better understanding of the potential consequences to removing inattentive respondents from their sample.

For this study, to control for measurement error of inattentive respondents a trap question was added to our conventional cheap talk script. The cheap talk script/trap question seen by respondents is displayed in Figure 1.3. Approximately half of the survey respondents were asked to answer the cheap talk script/trap question immediately after they finished the first portion of the choice experiment (either steak or ground beef questions). The other half of the respondents proceeded to the second portion of the choice experiment without being presented with the cheap talk script/trap question.

Statistical Model

Answers to the choice set questions can be modeled through Lancastrian consumer theory (Lancaster, 1966) as well as random utility theory (McFadden, 1973). Lancaster proposed consumers derive utility from individual product attributes and therefore, the total utility consumers derive from consumption of a product can be decomposed into separate utilities for their component characteristics. Random utility theory assumes consumers rationally choose the alternative from a given choice set that yields the highest level of utility. It follows from this theory that the probability of selecting a given alternative is higher if the utility provided by the alternative is the highest among the various other alternatives in the choice set. Given this theory, an individual i's utility given the j^{th} alternative is selected can be defined as

(1) $$U_{ij} = V_{ij} + \epsilon_{ij}$$

where U_{ij} is the utility obtained by individual i when selecting alternative j, V_{ij} is the systematic portion of the utility function and is determined by the attributes of the alternative j, and ϵ_{ij} is the stochastic component of the utility function not determined by the attributes. This stochastic component implies the true utility can never be observed. However, assuming the consumer is a utility maximizer, given a choice set (C_i) we know the consumer selects the alternative type *j*, that yields the highest utility among the possible alternatives in the choice set. The probability that alternative *j* is chosen can then be defined as

(2) $$Prob(j|C_i) = Prob\big(V_{ij} + \epsilon_{ij} \geq V_{ik} + \epsilon_{ik}; for\ all\ k \in C_i\big).$$

As demonstrated by McFadden (1983), as long as the stochastic errors in equation (1) are independently and identically distributed across the j alternatives and N individuals with a Type I Extreme Value distribution, the probability of individual *i* choosing alternative *j* is

(3) $$prob(j) = \frac{e^{V_{ij}}}{\sum_{k \in C} e^{V_{ik}}}.$$

Assuming linearity in the parameters of V_{ij} allows the functional form of $V_{ij}$to be expressed as

(4) $$V_{ij} = \boldsymbol{X}_{ij}'\boldsymbol{\beta}$$

where $\boldsymbol{X}_{ij}$ is a vector of attributes for alternative *j* and consumer *i*, and $\boldsymbol{\beta}$ is a vector of coefficients to be estimated. Equation (4) represents a standard conditional logit model.

Using the conditional logit model in equation 4 with price and the twelve discoloration indicator variables, the utility for steak and ground beef can be expressed as

(5) $$V_{ij} = \beta_1 Price_{ij} + \alpha_1 light25_{ij} + \alpha_2 light50_{ij} + \alpha_3 light75_{ij} + \alpha_4 light100_{ij} + \alpha_5 medium25_{ij} + \alpha_6 medium50_{ij} + \alpha_7 medium75_{ij} + \alpha_8 medium100_{ij} + \alpha_9 dark25_{ij} + \alpha_{10} dark50_{ij} + \alpha_{11} dark75_{ij} + \alpha_{12} dark100_{ij} + \varepsilon_{ij}$$

where V_{ij} is the utility of the ith individual having selected choice alternative j, $Price_{ij}$ is the price ($/lb.) of the choice alternative j chosen by individual i, ε_{ij} is the random error term following a Type I Extreme Value distribution, $light25_{ij}$ is a dummy variable equal to one if choice alternative j chosen by individual i had 25 percent coverage and light intensity of discoloration and equal to zero otherwise, and $light50_{ij}$ through $dark100_{ij}$ follow the same logic as the explanation for $light25_{ij}$.

Let (5) be written succinctly as

$$V_{ij} = \beta_1 Price_{ij} + \boldsymbol{X}_{ij}'\boldsymbol{\alpha} + \varepsilon_{ij} \tag{6}$$

where $\boldsymbol{X}_{ij}$ is a vector of indicator variables describing the product's coverage and intensity of discoloration and $\boldsymbol{\alpha}$ is a parameter vector.

As mentioned earlier this survey made use of cheap talk script combined with a trap question as a way to reduces hypothetical bias and measurement error resulting from inattentive respondents. Reductions in hypothetical bias can be modeled as an increase in the marginal utility of money, whereby a higher value means that each dollar received in income adds more to utility. As their marginal utility of money increases, people's willingness to pay for any attribute decreases as they value their money more and are less willing to trade money for goods. Previous literature suggests that those exposed to the cheap talk script exhibit a higher utility for money, and it seems reasonable to assume that the more attentive they are to the script, the greater the impact of that script on their utility of money. The negative of the coefficient for the price variable can be thought of as an individual's marginal utility for money. That is, for every dollar/lb. increase/decrease in the price of a steak or package of ground beef (taking away an individual's income) we estimate the respondent's

utility will decrease/increase by the amount of this coefficient. Notice that, because there is no 'none' option in the choice sets we do not estimate the overall value of a non-discolored steak but rather the discounts assigned to discolored steaks relative to a non-discolored steak. This has implications for how changes to the measured marginal utility of money are interpreted. Suppose the total willingness to pay for a whole non-discolored steak is WTP_{ND}, which can be decomposed into the utility of the steak (V_{ND}) divided by the marginal utility of money: $\left(WTP_{ND} = {V_{ND}}\Big/{-\beta_{price}}\right)$. This study does not estimate V_{ND} but it does estimate β_{price} as well as the change in utility from that steak becoming discolored. Let $V_{25\%,Light}$ be the utility of a steak with 25 percent coverage and light intensity of discoloration. This study is able to estimate $\Delta V_{25\%,Light} = V_{ND} - V_{25\%,Light}$, which should be negative if consumers dislike discoloration. The WTP premium/discount consumers apply to the *25%-Light* steak relative to the non-discolored steak is then $\left({\Delta V_{25\%,Light}}\Big/{-\beta_{price}}\right)$. This is a negative value expressed in dollars, making it a discount. The total value of the discolored steak is then

$$WTP_D = {V_{ND}}\Big/{-\beta_{price}} + {\Delta V_{25\%,Light}}\Big/{-\beta_{price}}. \quad (7)$$

There are two components to (7). One is the overall willingness to pay for non-discolored steaks, and the other is the discount assigned to the discolored streaks. When presented with a cheap talk script subjects presumably answer the questions more thoughtfully, attempting to give a more realistic depiction of their choices if, hypothetically, real money was used. It could be argued that consumers are more familiar with non-discolored beef than discolored beef. Some consumers may have never purchased discolored beef before, not even at a steep discount, largely because discolored steak is not displayed on store shelves as frequently as non-discolored steaks. Studies have shown hypothetical bias is

larger for unfamiliar goods than it is for familiar goods (Schläpfer and Fischhoff, 2012) which means that subjects may correct for hypothetical bias differently for the non-discolored and the discolored steaks. For example, subjects may hypothetically say they would pay $10 in total for a non-discolored steak and $7 for a discolored steak. However, when asked to take their purchases seriously, and behave as if they are spending real money, they might

- reduce the total amount they are willing to pay for each steak by the same amount, maintaining the $3 discount applied to the discolored steak,
- keep the total amount they are willing to pay for the non-discolored steak the same, as it is a familiar good, but reduce the amount they would pay for the discolored steak, thereby increasing the size of the discount applied to the discolored steak,
- reduce the total amount they are willing to pay for both steaks, but increase the size of the discount applied to the discolored steak, or
- reduce the total amount they are willing to pay for both steaks, but decrease the size of the discount applied to the discolored steak.

For example, being familiar with purchases of non-discolored steaks respondents may change their willingness to pay for those steaks only slightly, but when they take discolored steaks more seriously, being an unfamiliar item, they might make larger changes to the discounts assigned to those steaks.

Increasing the magnitude of β_{price} in (7), as would be expected when reducing hypothetical bias, would reduce the total WTP of both steaks, but it would also reduce the discount applied to the non-discolored steak, which may not be appropriate. In order to allow

both the overall WTP of the steaks to fall and the discount applied to non-discolored steaks to be negative, zero, or positive, the term $\Delta V_{25\%,Light}$ must be allowed to change after the adjustment for hypothetical bias, in addition to allowing β_{price} to change. This is achieved by expressing the utility function as

$$(8) \qquad V_{ij} = \beta_1 Price_{ij} + \beta_2 No\_CT_{ij} * Price_{ij} + \beta_3 TQF_{ij} * Price_{ij} + \boldsymbol{X}_{ij}'\boldsymbol{\alpha} + No\_CT_{ij} * \boldsymbol{X}_{ij}'\boldsymbol{\gamma} + TQF_{ij} * \boldsymbol{X}_{ij}'\boldsymbol{\eta} + \varepsilon_{ij}$$

where $\boldsymbol{\alpha}, \boldsymbol{\gamma},$ and, $\boldsymbol{\eta}$ are parameter vectors, No_CT_{ij} is a dummy variable equal to one if individual i was not been presented with the cheap talk/trap question when selecting alternative j and equal to zero otherwise (individuals where No_CT_{ij} = 1 are referred to as 'unexposed' subjects), TQF_{ij} is a dummy variable equal to one if individual i was presented with, and failed, the cheap talk/trap question (individuals where TQF_{ij} = 1 are referred to as 'exposed-inattentive' subjects). Thus an individual was presented with the cheap talk/trap question and passed it if $No\_CT_{ij} = TQF_{ij} = 0$; such individuals are referred to as 'exposed-attentive' subjects.

Results

Conditional Logit Estimation

For choice experiments to be reliably predictive, previous literature demonstrates the need for consumers to be knowledgeable about the product involved in the experiment (Beshears et al., 2008). For this reason, those respondents who indicate they have never purchased beef steak or ground beef are excluded in the respective steak and ground beef

models. These exclusions limit the total sample size for the steak model to 2,317 and for the ground beef model to 2,395.

Using equation (8) the results of the steak and ground beef models are summarized in Table 1.2. Consider the signs and significance of the discoloration variables (25%-Light to 100%-Dark). The case where all of these variables equal zero refers to a steak with no discoloration, so it is not surprising that all the coefficients are both negative and statistically significant ($p<0.001$), indicating consumers dislike discoloration of any kind. For both beef types, an increase in the coverage of discoloration while holding the intensity of discoloration constant usually results in a smaller number (larger in absolute value), indicating a greater disutility due to coverage of discoloration. The same can be said for an increase in intensity while holding the coverage constant. This is expected. However, sometimes the coefficients are counter-intuitive. For example, in the steak case the coefficient for 25%-Medium is actually larger (less negative) than the coefficient for 25%-Light, whereas one would expect greater coverage of discoloration would result in greater disutility. However, the null hypothesis for the likelihood ratio test that these coefficients are equal has a p-value of 0.389 indicating they are not statistically different. The same can be said for the coefficients for 75%-Dark and 100%-Dark; their values appear counter-intuitive but in reality are not statistically different (p-value 0.722).

There are also cases where the relationship between coefficients is as expected but are not statistically different. For example, one would expect the coefficient for 100%-Medium to be smaller than the coefficient for 75%-Medium for steaks, but likelihood ratio tests confirm that the coefficients are not statistically different (p-value 0.525). When such likelihood-ratio tests are conducted for greater differences in discoloration (*e.g.*, 25%-Light

and 100%-Dark) the null hypothesis that they are the same is usually rejected. This implies that overall discoloration does reduce the appeal of beef products, and the greater the discoloration, the greater the reduction, but relatively small changes in coverage and intensity of discoloration sometimes do not alter utility in a statistically significant manner.

Now consider how the cheap talk/trap questions influence response patterns across the three groups: (1) the unexposed (those who did not see the cheap talk script) (2) the exposed-attentive (those who saw it and passed the trap question) and (3) the exposed-inattentive (those who saw it and failed the trap question). Likelihood-ratio tests reject the null hypothesis that all of the coefficients for the interaction terms (in $\gamma, and\ \eta$) are zero, implying that the three groups demonstrate different response patterns. For both the steak and ground beef models the price coefficient is negative and significant at the one percent level. This indicates as the price of a beef item increases the utility from purchasing it decreases. To investigate the impact of the cheap talk/trap question, consider how the marginal utility of money (negative of the price coefficient) changes across the three groups. In the case of ground beef the marginal utility of money for the unexposed group is 0.2416 - 0.0272 = 0.2144; for the exposed-attentive group, 0.2416; and for the exposed-inattentive, 0.2416 - 0.0358 = 0.2058, which is a similar value to the unexposed group. The utility of money for the exposed-attentive group is about 15 percent larger than the other two groups, which would suggest they have a lower WTP for non-discolored steak. This is the expected result if the cheap talk script is effective. However, these three numbers are not statistically different, as indicated by the insignificant coefficients for $No\_CT_{ij} * Price_{ij}$ and $TQF_{ij} * Price_{ij}$. For the steak case the exposed-attentive group has a lower utility of money than the unexposed group, contrary to theory yet occasionally documented in other studies (Morrison and Brown,

2009), but again the differences are not statistically significant. Thus, overall, the cheap talk/trap question does not statistically alter the marginal utility of money for the participants.

However, as mentioned previously, this only implies that the value of the non-discolored steak is the same across the three groups. The discounts assigned to discolored steaks might vary between the groups, and since the discounts are a function of multiple parameters, changes in discounts across the groups are studied in the context of willingness-to-pay (WTP).

Willingness to Pay (WTP)

In a conditional logit model, coefficients cannot be directly interpreted as the direct effects of the explanatory variables on the probability of choosing each alternative (steak or ground beef). Rather, they represent the direct effects associated with each of the explanatory variables on the utility function, which can be used to calculate the WTP for each of the attributes. Estimates of WTP are calculated using the ratio of the coefficient of the attribute of interest over the marginal utility of money (negative of the price coefficient) as in $\left(-\frac{\beta_{attribute}}{\beta_{price}}\right)$. For example, the ratio $\left(-\frac{\beta_{Light100\%}}{\beta_{price}}\right)$ is the discount assigned to steaks with 100 percent coverage with light intensity of discoloration relative to a non-discolored steak. This value should be negative because the coefficients $\beta_{Light100\%}$ and β_{price} are both negative (if consumers dislike coloration and dislike paying money). If, for example, its value is -$2.00 it means that the maximum price consumers will pay for the 100%-Light discolored steak is $2.00/lb. less than the maximum price they will pay for the non-discolored steak.

Note that the probability of a person (in the exposed-attentive group) choosing a discolored steak over a non-discolored steak is calculated as

(9) $$[\exp(\beta_1 \Delta Price_{dis} + \boldsymbol{X}_{dis}'\boldsymbol{\alpha})]/[\exp(0) + \exp(\beta_1 \Delta Price_{dis} + \boldsymbol{X}_{dis}'\boldsymbol{\alpha})]$$

where $\Delta Price_{dis}$ is the price of the discolored steak minus the price of the non-discolored steak, $\boldsymbol{X}_{dis}$ is a vector of indicator variables describing the product's coverage and intensity of discoloration and $\boldsymbol{\alpha}$ is a parameter vector. The WTP discount associated with the discolored steak is the value for $\Delta Price_{dis}$ that sets this probability to 0.5. While WTP is often referred to as consumers' WTP, as if all consumers had homogenous preferences, in reality it measures the WTP for the average consumer. For example, if a discount of $2.00/lb. for a discolored steak is calculated, it implies that at a $2.00/lb. discount, if consumers are presented with one non-discolored and one discolored steak, half of consumers would purchase the discolored steak and half would purchase the non-discolored steak, making the 'average' consumer indifferent between the two and giving them a 50 percent chance of purchasing the discolored or non-discolored steak.

As mentioned previously, allowing the cheap talk/trap question to influence WTP only through a larger utility of money would guarantee that the discounts associated with discolored steaks would fall, while it seems plausible that consumers could actually increase the discounts. Thus, we allow both the numerator and the denominator to vary between the three groups. For example, while the discount for the exposed-attentive group for Light-100% discoloration is calculated as $\left(-\frac{\beta_{Light100\%}}{\beta_{price}}\right)$, for the unexposed group this numerator is adjusted by adding the coefficient for $No\_CT_{ij} * light100_{ij}$, and the denominator is adjusted by adding the term -$No\_CT_{ij} * Price_{ij}$.

The WTP discounts for each form of discoloration for steak and ground beef are reported in Table 1.3 and Table 1.4.

The WTP numbers are all negative because they refer not to the value of a steak but incremental amounts of discoloration. That is, they measure consumers' willingness to pay for greater discoloration, and since consumers do not want discoloration, the WTPs are negative. For the steak model a value of -$5.66 for the 25%-Light attribute in the exposed-attentive group can be interpreted to mean: for a steak that has 25 percent coverage and light discoloration intensity we would expect the average consumer to be willing to pay $5.66/lb. less than they would for the perfect (no discoloration) steak. Similarly, a value of -$4.49 for the 25%-Light attribute in the ground beef model can be interpreted to mean that for ground beef that has 25 percent coverage and light discoloration intensity we would expect the average consumer to be willing to pay $4.49/lb. less than they would for the perfect ground beef

To compare the WTP values across the three groups, we follow the Krinsky Robb method (parametric bootstrap). Within Tables 1.3 and 1.4, WTP values across the three groups are compared for each discoloration attribute. Those values within a row that are statistically different from one another differ in superscript letters. This allows us to quickly evaluate the effect of the cheap talk script as well as the trap question. One general trend is that when holding coverage constant at 25 percent the exposed-attentive and exposed-inattentive WTP values differ significantly ($p<.05$) for all levels of intensity within the steak model and for all but light intensity within the ground beef model. At higher levels of coverage this does not appear to be true. A plausible explanation for this would be that at higher levels of discoloration coverage, no matter the attention level, even at a steep discount respondents are unwilling to purchase the discolored products. However, at low levels of discoloration coverage the attentive respondents respond carefully (while considering

hypothetical bias) and take time to consider their true preferences while the inattentive may demonstrate a greater propensity to simply reject discoloration in favor of the 'perfect' product no matter the price. This result supports the use of the cheap talk script and trap question within the survey and demonstrates the divergence in results when the attention level of respondents is considered.

Discussion

Failing the Trap Model

As discussed previously, Berinsky, Margolis, and Sances (2014) recommend analyzing the predictors of trap question passage/failure. By following this recommendation researchers learn what the demographic predictors of trap passage are for the specific sample, which allows researchers to have a better understanding of how removal of inattentive respondents from the sample may induce bias.

Demographic information is collected through the survey including, gender, income, age, and education level. To test if these variables are proportionally different between those respondents who pass as compared to those who fail the trap question a Chi-squared test is conducted for each. Consider the test for gender, where "Female" is a dummy variable equal to one if the respondent is female and zero if male. The Chi-square test is used to determine if the proportion of females among the exposed-inattentive differs from the exposed-attentive group, and as Table 1.5 shows, this proportion is not significantly different at the 0.05 level. We thus conclude that in terms of gender, using the WTP for the exposed-attentive group only does not bias the results. Similar tests are conducted for "Education", "Age", and "Income", which are all categorical variables. As with gender there are no income or education differences between the exposed-attentive and the exposed-inattentive subjects.

However, there are statistically significant differences in the ages of the two groups, and a correlation analysis reveals that older subjects are less likely to fail the trap test. These results are summarized in Table 1.5.

These results help us get a better picture of what type of demographics are contributing to trap failure. Because the chi-squared test for 'age' indicates significant differences, it is evident those respondents who fail the trap test are not distributed randomly throughout the entire sample. Looking at the distribution of respondent ages within the group who passed the trap question reveals that while this subset of the sample still contains a diverse distribution of ages, the median age increases as compared to the entire sample (which fit the U.S. population reasonably well) to approximately 50 years old. If the responses of the exposed-inattentive group are simply discarded, as is often the common practice, then the survey sample data may be misrepresentative of the desired population (U.S. consumers). For this reason, we follow the recommendation of Berinsky, Margolis, and Sances (2014) and don't discard all who fail the trap but rather display the results stratified by level of attention; exposed-attentive, exposed-inattentive, and unexposed. This allows readers to easily see how the results change with attention level and promotes transparency in the presentation of results.

Strategic discounts for quick sale of discolored beef

Overall these discounts are large, especially relative to the listed prices of the beef products. The price of an average steak in 2017 was around $7.93/lb., while the price of ground beef was $3.71/lb. Let us assume the average prices of $7.93/lb. and $3.71/lb. for the steak and ground beef respectively represent the price of non-discolored ribeye steak and ground beef, respectively. To calculate the market price for a discolored steak for an average

consumer to be indifferent when comparing the discolored steak to a non-discolored steak we can add the average WTP discounts estimated for specific discoloration attributes to the prices of the 'perfect' steaks or ground beef packages. For example, adding the average WTP estimate of -$5.66/lb. for a steak that has 25 percent coverage and classified as having a light intensity of discoloration to the perfect steak price of $7.93/lb. results in an estimated market price of $2.27/lb. for a steak with those discoloration attributes.

Scanning over the WTP estimates calculated for steak displays estimates ranging in value from -$5.66/lb. (25%-Light) to -$20.14/lb. (75%-Dark). It is evident any steak that has at least 50 percent of its surface area discolored and classified as either light, medium, or dark intensity of discoloration would be expected to be unmarketable to the average consumer (when compared to the average 2017 price) as the negative WTP estimate for these discoloration attributes would exceed the market price of the perfect steak. The resulting estimated market prices for steaks with these discoloration attributes would be negative. This would essentially indicate the average consumer would have to be compensated before purchasing a steak with these discoloration attributes. In ground beef the smallest WTP discount estimated is -$4.13/lb. (25%-Medium) and the price of the perfect ground beef is only $3.71/lb. (2017 average price per lb. in the U.S.). This would indicate any ground beef that has 25 percent coverage and classified as having either light, medium, or dark intensity of discoloration would be expected to be unmarketable to the average consumer as the negative WTP estimate for these discoloration attributes would exceed the market price of the perfect ground beef.

This demonstrates that the average consumer displays a considerable aversion to discoloration. However, this does not indicate that we should expect retailers to begin paying

consumers to 'purchase' discolored beef products or even offer them at steep discounts (e.g. >50 percent). Three reasons motivate this claim. First, many consumers may have preferences that differ considerably from the average consumer. Second, stores may have only a few discolored items to sell relative to non-discolored beef items. Third, many stores have other options available to them other than simply discarding unsold raw beef.

Consider a store that has a number of steaks on display for sale, and all are basically identical except for one, and it differs only due to 25%-Light discoloration. With many identical non-discolored steaks and one discolored but otherwise identical steak the consumer problem reduces to a choice between a non-discolored and a discolored steak. This discolored steak requires a $5.66/lb. discount before the average (attentive) consumer will be indifferent between it and a non-discolored steak, as [exp{(-0.172)(-$5.66)+ (-0.9732)}]/[exp(0)+exp{(-0.172)(-$5.66)+(-0.9732)}]=0.5. However, suppose the discount offered is only $1.00/lb., making the probability of a consumer purchasing the discolored steak about 30 percent. Assume also that customers arrive at meat display at the rate of one every five minutes and each customer purchases one steak. The probability the store will sell the discolored steak (assuming there is always at least one non-discolored steak) in less than n*5 minutes can be calculated as $(1-(0.7)^n)$. The probability they will sell the discolored steak within thirty-five minutes is then greater than 90 percent. Thus, while the average consumer discounts discolored meats heavily, with enough shoppers the store can expect to sell a discolored steak within a reasonable amount of time at a much smaller discount.

Discolored beef products are sold in low quantities compared to non-discolored beef products. Thus, it is natural to assume that a retailer attempting to maximize profits will not

price discolored beef using the average WTP discount but rather a smaller discount, targeting the less discriminating consumer instead of the average consumer.

Additionally, many stores may have a section where they offer cooked foods, and it is likely they may use discolored beef that has not sold as an input into these cooked foods, such as making meatloaf from discolored ground beef. For this reason the store may only be willing to sell the discolored beef at a small discount, as the opportunity cost of not selling it as a cooked food item is high.

These results illustrate that even at low levels of discoloration the average consumer has little appetite (WTP) for discolored beef. This would indicate that not only is there a need for, but that retailers should be interested in, technology that would prolong the shelf life of the vibrancy of the bright, cherry-red color of fresh beef. Much research is being done in this field. Suman et al. (2014) provides a through overview of the current literature on strategies to improve beef color and stability of color. Their work highlights numerous noteworthy efforts within the meat industry using both pre-harvest (i.e. diet, animal management) and post-harvest (i.e. packaging, aging, antioxidants) intervention strategies to improve color stability in beef. They note that much of the current research suggests that the effects of several of these strategies are specific to type of animal, feeding regimen, packaging system, and muscle source. Their concluding recommendation is that "meat scientists should explore novel ways to manipulate these factors using a biosystems approach to achieve improved beef color stability, satisfy consumer perception, and increase market profitability" (Suman, et al. 2014, pp 490). Our results for this study only highlight the continued need for further research and advancement in this area.

Limitations

Though this study makes a concerted attempt to control for hypothetical bias it is important to note that without the use of revealed preference methodologies hypothetical bias is always a concern. Our results indicate a strong aversion to even relatively low levels of discoloration in beef. Within this hypothetical survey respondents entire responsibility is to focus on the choice experiment questions and make selections based off of the varying attributes of the meat products. This could result in participants focusing on the negative aspects of the discoloration attributes more than might otherwise mimic reality compared to the focus of shoppers in a retail environment. Shoppers typically go to a grocery store with the purpose of buying multiple products and are constrained by time. It is therefore possible that while in our study we find a strong aversion to discoloration, this same level of aversion may not be found if we could evaluate consumers actual purchase decisions. Given their limited time, the level of attention that consumers actually place on the color of the beef they purchase could result in different revealed preferences than those stated in our hypothetical survey.

As WTP estimates for varying levels of discoloration have yet to be studied, this research looks at discoloration attributes over the entire continuum of possibilities (0-100 percent coverage; light, medium, and dark intensity of discoloration) to begin to get an overall understanding of consumer preferences and WTP for discoloration attributes in beef. As the results indicate average consumers would require steep WTP discounts for discolored beef with even 25 percent of its surface area discolored, further research is needed to understand the WTP for discoloration within the range of 0-25 percent discoloration coverage. Additionally, this research selected steaks and ground beef as the beef products to

focus on. As research continues in this area, additional beef cuts should be used to better understand how consumer preferences for discoloration change within higher and lower quality beef cuts. Demartini et al. (2018) demonstrated that by grouping respondents together in clusters based on their attitudes and then adding this information to the model through additional interaction terms, niche market preferences may be better identified. As discolored beef is unfamiliar to many consumers, it is reasonable to expect consumer attitudes towards beef discoloration to vary considerably. Future research in this area may benefit through additional analysis including clustering of consumers by attitudes.

Conclusions

Research has shown consumers prefer beef to be bright, cherry-red in color. Deviations from this color are perceived by consumers as a possible indication of unwholesomeness. Often meat begins to lose its bright, cherry-red color long before it begins to spoil, resulting in meat that is either marked down in price to entice a quick sale, ground for other products, or in some cases discarded completely. This leads to waste and loss of revenue within the beef market. While consumers' avoidance of discolored meat is well-known there is little information regarding the extent to which discoloration reduces its value. This study demonstrates that even at low levels of discoloration the value of a beef item can fall as much as 50 percent.

Table 1.1. Summary Statistics of Survey Respondents

Description	Category	Percentage
Gender	Female	51.98%
	Male	48.02%
Age	18 to 25 years old	17.72%
	26 to 34 years old	15.41%
	35 to 49 years old†	23.72%
	50 to 65 years old	25.95%
	more than 65 years old	17.21%
Education	less than high school	1.25%
	high school/GED	18.31%
	some college	24.19%
	2-year college degree†	11.25%
	4-year college degree	29.95%
	graduate degree	15.05%
Income	less than $30,000	26.93%
	$30,000 to $59,000†	27.75%
	$60,000 to $89,999	19.33%
	$90,000 to $119,999	11.84%
	$120,000 to $149,999	7.53%
	$150,000 or more	6.62%
Shopper	Is the primary shopper in the household	88.51%
	Is not the primary shopper in the household	11.49%
Frequency of steak purchase	never	9.17%
	less than once a month	22.74%
	about once a month	23.48%
	about once every two weeks	22.89%
	about once a week	16.35%
	more than once a week	5.37%
Frequency of ground beef purchase	never	6.12%
	less than once a month	14.11%
	about once a month	20.85%
	about once every two weeks	27.52%
	about once a week	24.54%
	more than once a week	6.86%

Note: † Indicates median category.

Table 1.2. Steak and Ground Beef Model Results

	Steak			Ground Beef		
Variable	Parameter		Standard Error	Parameter		Standard Error
Price	-0.1720	***	0.020	-0.2416	***	0.029
No Cheap Talk x Price	-0.0230		0.022	0.0272		0.031
Trap Question Failed x Price	0.0208		0.024	0.0358		0.032
25%-Light	-0.9732	***	0.084	-1.0858	***	0.085
50%-Light	-1.9401	***	0.087	-2.2545	***	0.097
75%-Light	-3.2192	***	0.158	-2.5794	***	0.121
100%-Light	-3.2217	***	0.173	-3.9502	***	0.234
25%-Medium	-1.1536	***	0.081	-0.9975	***	0.080
50%-Medium	-1.6576	***	0.092	-1.7279	***	0.091
75%-Medium	-3.0370	***	0.150	-3.0205	***	0.154
100%-Medium	-3.1679	***	0.153	-4.0345	***	0.238
25%-Dark	-1.0768	***	0.087	-1.0370	***	0.085
50%-Dark	-3.2054	***	0.164	-3.0364	***	0.157
75%-Dark	-3.4649	***	0.176	-3.7176	***	0.203
100%-Dark	-3.2285	***	0.156	-3.6246	***	0.180
25%-Light x Trap Fail	-0.1920	*	0.100	-0.0211		0.098
50%-Light x Trap Fail	-0.0871		0.104	0.0518		0.115
75%-Light x Trap Fail	0.2835		0.179	0.2366	*	0.135
100%-Light x Trap Fail	0.3157	*	0.189	0.6249	***	0.241
25%-Medium x Trap Fail	-0.1700	*	0.096	-0.0623		0.092
50%-Medium x Trap Fail	-0.0751		0.108	-0.0305		0.106
75%-Medium x Trap Fail	0.2482		0.162	0.2987	*	0.172
100%-Medium x Trap Fail	0.5600	***	0.156	0.5646	**	0.267
25%-Dark x Trap Fail	-0.1530		0.103	-0.0759		0.099
50%-Dark x Trap Fail	0.3592	**	0.166	0.2739		0.170
75%-Dark x Trap Fail	0.5026	***	0.173	0.4339	**	0.212
100%-Dark x Trap Fail	0.3292	*	0.168	0.5347	***	0.192
25%-Light x No Cheap Talk	-0.0381		0.090	0.0110		0.090
50%-Light x No Cheap Talk	-0.0476		0.094	-0.2559	**	0.104
75%-Light x No Cheap Talk	-0.1208		0.169	-0.0424		0.129
100%-Light x No Cheap Talk	-0.0623		0.183	-0.0625		0.244
25%-Medium x No Cheap Talk	0.0283		0.087	-0.0334		0.085
50%-Medium x No Cheap Talk	-0.0336		0.099	-0.0637		0.097
75%-Medium x No Cheap Talk	0.1367		0.158	-0.1493		0.164
100%-Medium x No Cheap Talk	0.1119		0.159	-0.5085	**	0.258
25%-Dark x No Cheap Talk	-0.0504		0.093	-0.1122		0.091
50%-Dark x No Cheap Talk	0.2803	*	0.170	-0.0073		0.166
75%-Dark x No Cheap Talk	0.3303	*	0.182	0.0233		0.212
100%-Dark x No Cheap Talk	0.1261		0.164	-0.0577		0.190

***, **, * Statistically significant at the 0.01, 0.05, and 0.1 level
Number of observations: Steak= 30,121; Ground Beef= 31,135
Log Likelihood: Steak= -20,677; Ground Beef = -19,415

Table 1.3. Willingness-to-Pay Estimates for Discoloration Attributes in Steak

Attribute	Exposed Attentive	Exposed Inattentive	Non-Exposed
25%-Light	-$5.66[a]	-$7.71[b]	-$5.19[a]
50%-Light	-$11.28[ab]	-$13.41[b]	-$10.19[a]
75%-Light	-$18.72[a]	-$19.42[a]	-$17.13[a]
100%-Light	-$18.73[a]	-$19.22[a]	-$16.84[a]
25%-Medium	-$6.71[a]	-$8.75[b]	-$5.77[a]
50%-Medium	-$9.64[ab]	-$11.46[b]	-$8.67[a]
75%-Medium	-$17.66[a]	-$18.44[a]	-$14.87[a]
100%-Medium	-$18.42[a]	-$17.25[a]	-$15.67[a]
25%-Dark	-$6.26[a]	-$8.13[b]	-$5.78[a]
50%-Dark	-$18.64[a]	-$18.82[ab]	-$15.00[b]
75%-Dark	-$20.14[a]	-$19.59[ab]	-$16.07[b]
100%-Dark	-$18.77[a]	-$19.18[a]	-$15.91[a]

Note: WTP estimates within a row followed by different letters differ significantly ($p<0.05$).

Table 1.4. Willingness-to-Pay Estimates for Discoloration Attributes in Ground Beef

Attribute	Exposed Attentive	Exposed Inattentive	Non-Exposed
25%-Light	-$4.49[a]	-$5.38[a]	-$5.01[a]
50%-Light	-$9.33[a]	-$10.70[ab]	-$11.71[b]
75%-Light	-$10.68[a]	-$11.38[a]	-$12.23[a]
100%-Light	-$16.35[a]	-$16.16[a]	-$18.72[a]
25%-Medium	-$4.13[a]	-$5.15[b]	-$4.81[ab]
50%-Medium	-$7.15[a]	-$8.54[a]	-$8.36[a]
75%-Medium	-$12.50[a]	-$13.23[a]	-$14.78[a]
100%-Medium	-$16.70[a]	-$16.86[ab]	-$21.19[b]
25%-Dark	-$4.29[a]	-$5.41[b]	-$5.36[b]
50%-Dark	-$12.57[a]	-$13.42[a]	-$14.20[a]
75%-Dark	-$15.39[a]	-$15.96[a]	-$17.23[a]
100%-Dark	-$15.00[a]	-$15.01[a]	-$17.17[a]

Note: WTP estimates within a row followed by different letters differ significantly ($p<0.05$).

Table 1.5. Comparing Demographics of Respondents Who Passed vs. Failed the Trap

Variable	Chi-Square Test Statistic	P-value	Interpretation
Income	8.4182	0.1346	Not significant
Age	51.0798	<0.001	Older respondents are less likely to fail the trap
Education	10.7533	0.0565	Not significant
Female	0.1662	0.6835	Not significant

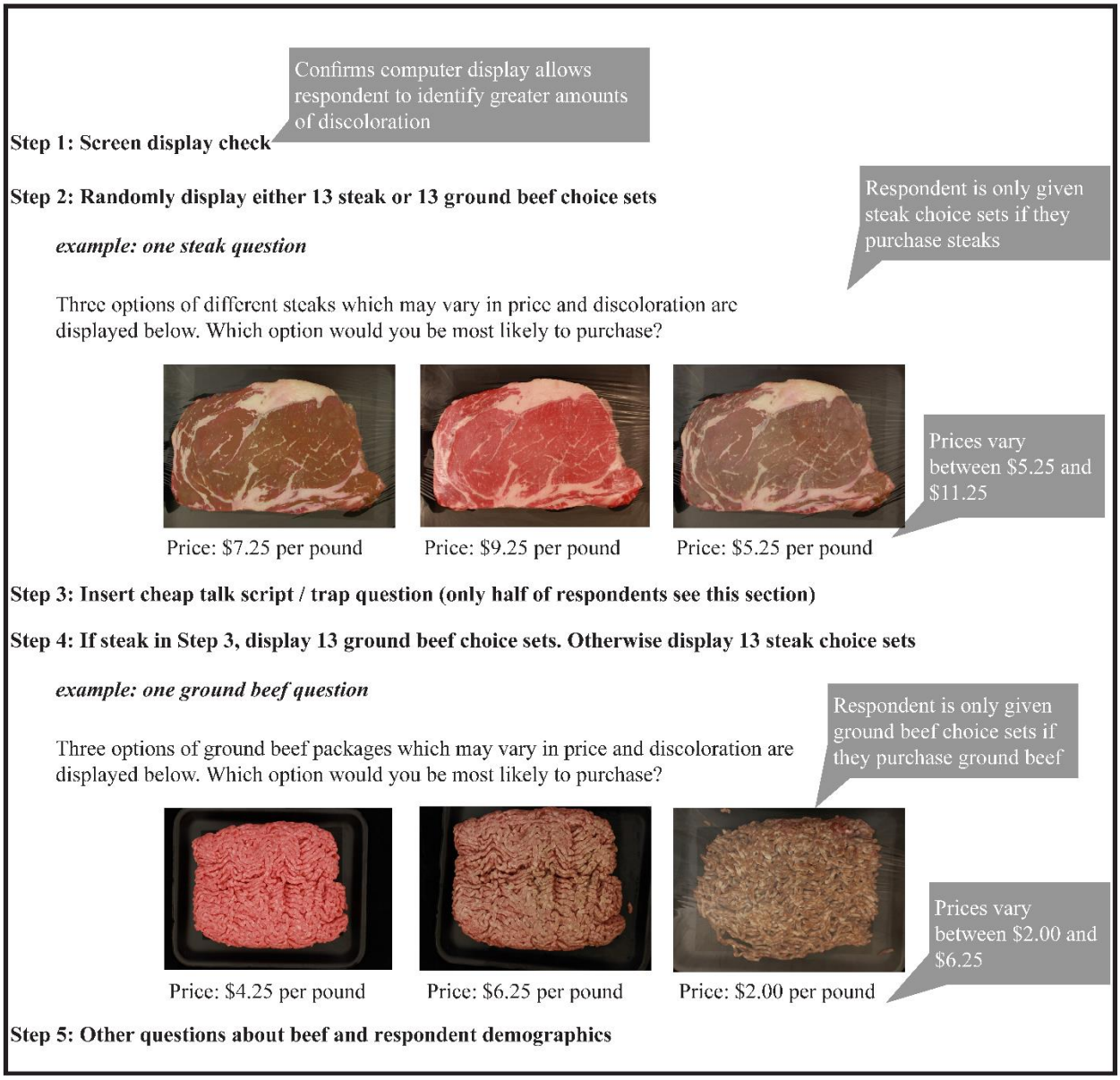


Figure 1.1. Overview of Survey

Figure 1.2. Twelve levels of discoloration used in the choice experiment

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 beef products you are evaluating. This purchase was hypothetical (as it is 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 in a store, i.e., noting that buying a product means that you would have less money available for other purchases. Of course, providing this information only matters if survey respondents read it. To demonstrate that you have read this information, please select the option that says it was left blank intentionally, instead of the option 'I have read the above information'.

○ Yes, I have read the above information

○ This option intentionally left blank

Figure 1.3. Cheap Talk Script/Trap Question

CHAPTER II

THE SPILLOVER EFFECT OF MARKETING DISCOLORED BEEF ON CONSUMER PREFERENCES FOR NON-DISCOLORED BEEF

Abstract

Consumers prefer retail beef to be bright, cherry-red in color. Retailers often mark down the price of discolored beef for quick sale. However, following this practice could result in a net loss of revenue if consumer willingness-to-pay (WTP) for non-discolored beef is negatively affected by the presence of discolored beef in the consumer choice set. Through a hypothetical online survey as well as a controlled in-person experiment we determine that for the majority of consumers, marketing discolored beef together with non-discolored beef increases consumers' evaluation of, but not their WTP for, non-discolored beef.

Keywords: beef color; beef discoloration; finite mixture model; willingness-to-pay

Introduction

In the retail beef market, past research has focused extensively on intrinsic attributes of meat such as fat content/marbling, tenderness, and taste (Lusk et al., 2001; Feuz et al., 2004; Lusk and Parker, 2009), as well as on less tangible attributes such as geographical origin of production and traceability (Loureiro and Umberger, 2003), method of production such as conventional versus organic, hormone-free, GM-free (Tonsor et al., 2005), and additional labeling preferences indicating such qualities as safety (Lim et al., 2013).

Among these attributes, the color of the meat can influence consumer decisions. Quality inferences are based in part on the color of the meat. Research has shown the preferred color of beef is bright cherry-red (Carpenter, Cornforth, and Whittier 2001; Killinger et al. 2004; Ramanathan and Mancini 2018). Through prolonged exposure to oxygen retail beef begins to discolor, turning brownish-red to brown. Consumers often perceive this discoloration as a condition of "unwholesomeness" (Faustman and Cassens, 1989). This decreases the willingness-to-pay (WTP) for the discolored beef (Grebitus, Jensen, and Roosen, 2013; Grebitus, Jensen, Roosen, and Sebranek, 2013;). As a result, beef cuts must be marked down in price, faced and repackaged, ground, or discarded (Smith et al., 1996). Of these options, it is unclear which is the most popular with retailers. It is, however, not unusual to find discolored beef cuts routinely marked down for quick sale at many grocery outlets. Retailers recognize consumers have a lower WTP for discolored beef yet hope to minimize loss of revenue by marking down the price instead of simply discarding. However, perhaps this marketing decision has been oversimplified. Discolored beef marked down in price and sold does provide additional

revenue that would otherwise be lost if the meat were discarded. However, what effect does marketing this discolored beef alongside non-discolored (bright, cherry-red) beef have on the value of the non-discolored beef? That is, is there a spillover effect, whereby the presence of a discolored steak alters the perceived quality of non-discolored steaks?

It is conceivable that the spillover effect is negative, meaning the value of the non-discolored steak falls due to the presence of a nearby discolored steak. If consumers believe it is poor meat management practices that causes the discoloration, they might believe that those poor practices threaten the safety of all the meat in the store.

Conversely, the presence of a discolored steak might raise the perceived quality of the non-discolored steak by making it seem more appealing by comparison. This would be a positive spillover effect. Also possible, would be that marketing discolored steaks alongside non-discolored steaks may have no effect on consumer preferences for the non-discolored steaks. Or, perhaps different subsets of consumers display all three reactions, with some liking the non-discolored steak more, some less, and some the same. The objective of this research is to measure the direction and size of the spillover effect both for consumers overall and subsets of consumers

The objective is achieved using two sets of data. One comes from an internet survey of Americans where they are directly asked how the presence of a discolored steak alters their perception of a non-discolored steak. They are also administered an open-ended question regarding their perception on what causes discoloration, as the perceived reasons for discoloration would influence the direction of the spillover effect. The second set of data comes from a controlled experiment where subjects view three steaks in a choice set and rate the appeal of the most and least appealing steaks. By randomly

varying whether a discolored steak is present in the choice sets we are able to measure the spillover effect. Throughout the remainder of this paper we refer to the internet survey as '*Study 1'* while the controlled, in-person experiment is referred to as '*Study 2'.*

Literature on spillover effects

While some economic models might explain the value of a good solely in terms of its attributes, a wealth of studies have demonstrated how the attributes of other goods in a choice set can influence that good's value. The marketing literature has devoted many resources to studying the 'decoy' effect, also referred to as an 'attraction' effect, whereby the addition of a third option alters a consumer's preference for the two original items (Heath and Chatterjee, 1995). Specifically, suppose there are two items, A and B. Some consumers prefer A and some consumers prefer B. If a third product is now added that is dominated by A but not B across all attributes (say, both in terms of price and quality) then consumers increase their purchases of A. Likewise if the third product is dominated by B but not A then its presence will increase purchases of B.

To illustrate, suppose there are originally two steaks for sale, one of higher quality (ribeye) but also a higher per pound price and the other of lower quality (sirloin) and at a lower price. If a discolored ribeye is then added to the choice set for sale at the same price as the non-discolored ribeye its presence could increase sales of the non-discolored ribeye.

Other studies have documented a 'compromise' effect, where consumers increase their perceived value of an item if it is depicted as a moderate option rather than an extreme option (Levav, Kivetz, and Cho, 2010). In this case consumers might state their WTP for a low-quality sirloin steak as $5.00 per lb., but then increase this WTP when a

discolored sirloin steak is added to the choice set. While the steak originally seemed 'low' quality the presence of the discolored steak now makes it seem a more 'moderate' quality, thereby increasing its value to consumers.

Both the attraction and compromise effects are more likely to occur when consumers have less familiarity with the good and its attributes (Levav, Kivetz, and Cho, 2010). The extent to which the attributes are 'meaningful' also influences decoy effects. If steak grades like Select, Choice, and Prime are less meaningful to a consumer, meaning they have difficulty understanding the differences in the grades, they are more likely to display the decoy effect (Ratneshwar, Shocker, and Steward, 1987), though this might be defined by some researchers as less familiarity.

The value of a good can be dependent upon the other goods in a choice set in cases where consumers exhibit bounded rationality and must employ cognitive effort and heuristics to arrive at a choice. That is, consumers are not just choosing the good that they prefer the most, they are actively trying to decide which one they prefer the most from a choice set and they are making this decision partly from environmental cues. The presence of a discolored steak in a choice set might provide cues the individual uses to infer the quality of other non-discolored steaks. If a person associates discoloration with poor management they might infer that the non-discolored steaks, though appearing safe to eat, were also managed poorly and could thus pose safety hazards. This would cause the person to reduce the value they place on non-discolored steaks in the presence of discolored steaks.

There is reasonable evidence to suggest that the presence of discolored steaks affects the value consumers place on non-discolored steaks. The preceding paragraph

suggests that a discolored steak reduces the value of its non-discolored neighbors, while the decoy and compromise effects suggest it would increase it. Or, the effects could offset one another such that a discolored steak has no impact on surrounding non-discolored steaks.

Methods and Data

Study 1

Is there a spillover effect, whereby the presence of a discolored steak makes non-discolored steaks more or less appealing? One way of answering this is to simply pose the question to consumers. In the fall of 2017 an internet survey was administered to 2,598 US respondents. This survey was administered on the Qualtrics survey platform (Qualtrics, 2018) and the sample was acquired by Survey Sampling International, LLC (2017). The sample was intended to be representative of all US consumers and overall the sample demographics fit the desired population reasonably well. A summary of the sample demographics can be seen in Appendix 2A. Additional discussion on how well the survey sample fit the targeted demographics is also included in Appendix 2A.

Among the questions asked was one concerning the spillover effect. Subjects were shown a picture of two steaks, one discolored, and the other non-discolored, and were explicitly asked how the discolored steak affects the perceived quality of the non-discolored steak. Figure 2.1 contains the question as it was asked in the survey, though the order of the questions was randomized across respondents. By analyzing how responses to the two questions differ across respondents we can begin to get an idea of what the spillover effect may be.

The survey also posed the open-ended question, "What do you believe causes discoloration in raw beef? Please answer in your own words using at least one sentence." By analyzing the content of the answers we can assess whether people believe discoloration is caused by a factor that could affect all other beef products or whether it is caused by a factor unique to each individual beef item, and thereby determine whether the reasons for discoloration should suggest the presence of a spillover effect. Additionally, respondents were asked what they do with discolored meat at home and whether they believe grocery stores should sell discolored meat, which aids in understanding whether discoloration is seen as a systemic management problem that would affect the value of all meats sold in the store.

Study 2

The second study consisted of a controlled in-person experiment where subjects were asked to rate the appeal of a set of steaks, with some participant choice sets containing a discolored steak and others only containing non-discolored steaks. The experiment was conducted on March 12, 2018 at the Robert M. Kerr Food and Agricultural Products Center on the Oklahoma State University campus. In total 118 people participated in the experiment though one subject did not complete the entire questionnaire, resulting in 117 observations. Subjects were recruited by sending out email invitations to the Division of Agricultural Sciences and Natural Resources at Oklahoma State University. Additionally, a sign was placed on the sidewalk adjacent to the building where the experiment was held inviting people to participate. Key demographic variables of participants were collected and summary statistics for the participant demographics are displayed in Table 2.1.

In 2017 the U.S. population was estimated as 50.8% female with a median income of $55,322 and a median age of 37.7 years (United State Census Bureau 2018). In comparison our sample has a higher percentage female, with a lower median income, and lower median age. This comes as a result of the experiment being held on a university campus where those choosing to participate in the experiment tend to be of a lower age and lesser income. Note, however, that most of the subjects are not college students, and women are more likely to shop for the family than men, and so in some aspects our sample better reflects the average shopper than a perfectly representative sample (Flagg et al., 2014).

Three Choice grade ribeye steaks (pH = 5.6), approximately 3.5-inch each, were purchased from a local meat purveyor. Four days before the experiment was held the steaks were cut horizontally through the center to yield a total of six steaks to be used in the experiment. By cutting the steaks in this fashion each steak could be paired with a nearly identical or mirror image counterpart steak. This in turn resulted in less variability between the steaks and helped to eliminate confounding variables within the experiment. Three of the non-paired steaks were promptly vacuum sealed and placed in a dark cooler until the day of the experiment. The other three steaks were placed onto foam trays with absorbent pads and over-wrapped with a PVC film (oxygen-permeable polyvinyl chloride fresh meat film). After packaging, the steaks were placed in a coffin-style open display case maintained at 2°C ± 1 under continuous lighting for four days. During these four days the steaks became increasingly discolored, and on the fourth day (day of the experiment) the coloring was measured using a HunterLab MiniScan® XE Plus spectrophotometer (Model 45/0 LAV, 2.5-cm diameter aperture, illuminant A, 10°

observer; HunterLab, Reston, VA). Both reflectance spectra from 400 to 700 nm (10 nm increments) and CIE a* values were measured on each steak at three random locations and the subsamples were averaged. CIE a* values indicate redness. A greater value indicates brighter red color, while a lower value indicates less red color. Discoloration or percent metmyoglobin is calculated using AMSA (2012) equations. Metmyoglobin is the pigment form that gives brown color. A greater number indicates more discoloration and a lower number indicates less discoloration. The discoloration is also documented using a digital Cannon SLR camera. Summaries of the mean CIE a* values as well as the mean percent metmyoglobin values for both the discolored and non-discolored steaks used in study 2 are displayed in Table 2.2.

On the day of the experiment, approximately 30 minutes before participants arrived (just long enough for the fresh beef to bloom and assume a bright, cherry-red color) the vacuum sealed steaks were removed from the dark cooler and packaged in a similar fashion to the discolored steaks on foam trays under PVC overwrap. As the time for the experiment to be held arrived we then had three non-discolored and three discolored steaks to be used in the treatments.

There were two treatments for the experiment with approximately half of the participants randomly receiving the first treatment while the other half received the second treatment. In the *discolored* treatment, three steaks were presented in a retail style cooler all on separate foam trays placed under clear PVC overwrap as would be found at a retail store ready for purchase. Participants were escorted to the retail cooler (two at a time were allowed to view the steaks). Two of the steaks presented were non-discolored steaks while the other was discolored, all three of which were randomly chosen and

varied for each set of two participants in the experiment. In the *non-discolored* treatment all three steaks were non-discolored. Subjects were randomly assigned to the discolored / non-discolored treatments by regularly rotating the treatments throughout the length of the experiment. The order the steaks were displayed was also randomized. Each participant was initially given "Survey Sheet 1" to complete. At the top of the first page of this survey sheet the participants were instructed to first, carefully observe the three steaks in front of them. Then they were asked to pick the steak they liked the most/least (randomly assigned to consider most or least) and with their chosen steak in mind, indicate to what extent they agreed or disagreed with the statements that followed. While the subjects were not asked to indicate which steaks were their most and least favorite, it is assumed (since all steaks are virtually identical except for color) that a discolored steak is never the most favored steak. The second page of "Survey Sheet 1" was identical to the first except that participants were instructed to respond to the statements while considering the opposite choice, most or least liked steak. The twelve statements (six statements repeated once) contained on "Survey Sheet 1" are as follows:

1. The quality of this ribeye steak is high.
2. I would seriously consider purchasing this steak.
3. If offered at a reasonable price, people would purchase this steak.
4. Most people would like the taste of this steak.
5. I would purchase this steak for ($X) per pound. (X varied randomly by $0.25 increments from $5.00 to $15.00)
6. I would purchase this steak for ($X) per pound. (X varied randomly by $0.25 increments from $5.00 to $15.00)

Each of the statements were followed by a Likert scale of agreement with seven choice options for the participants selection (strongly disagree, disagree, somewhat disagree, neither agree nor disagree, somewhat agree, agree, strongly agree). Appendix 2B shows "Survey Sheet 1" if the respondent was asked to appraise their least preferred steak on the first page.

After completing "Survey Sheet 1", the two participants were escorted away from the retail cooler and given "Survey Sheet 2" to complete. This portion of the survey collected the demographic information previously summarized in Table 2.1. Upon completion of "Survey Sheet 2", participants were given $10 in appreciation for their participation. The responses to the statements are summarized for both the steak participants liked the most as well as the one they liked the least in Table 2.3.

The questionnaire results are analyzed as follows. First a factor analysis is performed on all six statements to discern the number of latent factors that might be influencing response patterns, with the results shown in Table 2.4. This analysis is needed to determine what methodology is most appropriate for analyzing the data. For the most desirable steak the loadings are considerably higher for the first four statements than the last two. This is understandable, as the first four statements are simple overall appraisals of the steak whereas the last two statements also have the consumer evaluate the steaks relative to a random price. This additional information of price, and its random nature, ensures that unique factors will have a greater impact on their responses, reducing the extent to which the latent factor explains their responses.

For the least desired steaks the loadings are high for all six statements, suggesting just one latent construct (appraisal of the steak) performs well at explaining all responses. The loadings for the last two statements do not fall as much as they do for the most desired steak. This is perhaps because the least desired steak is often discolored, and consumers pay less attention to its price.

While the factor analysis finds a single factor explains responses to all statements, suggesting the responses to the Likert scales for all six statements can be summed to create one variable denoting steak desirability, we choose to combine statements 1-4 while separately combining statements 5 and 6 for analysis. We do this for three reasons. First, the sizes of the factor loadings do differ between these sets of statements, and second, the statements are of a different nature. Third, by analyzing the combined statements 5 and 6 separately we are able to estimate WTP for each steak type (preferred most or preferred least).

The objective of study 2 is to determine how the presence of a discolored steak will influence the appeal and value of the non-discolored steak, which is assumed to be the most desired steak (in all treatments there are at least two non-discolored steaks). The response patterns to the first four statements are analyzed by first summing the scores from (1 = strongly disagree to 7 = strongly agree) and analyzing how this sum changes when a discolored steak is included in a choice set, using a finite mixture regression model. The scores can be summed to create one variable because there is only one latent factor influencing their value. This model can be represented as

(1) $$S_{i,c} = \alpha_{0,c} + \alpha_{1,c}(D_i) + X_i\beta_c + e_{i,c}$$

where $S_{i,c}$ is the sum of the response to the four statements for the steak participants liked most, D_i is an indicator variable that equals one if a discolored steak is present and zero otherwise, X_i is a vector of demographic variables, $e_{i,c}$ is an error term, and $\alpha_{0,c}$, $\alpha_{1,c}$ and β_c are parameters to be estimated. The subscript "c" refers to classes of individuals. If a regular regression is used then all individuals belong to one class, c = 1 for everyone, and all respondents share the same parameter estimates. However, when a finite mixture model is used two classes of respondents are allowed (more were considered but models failed to converge beyond two classes), c = 1 or 2, and each class possesses its own distinct parameter estimates.

The spillover effect depends on the value and statistical significance of the $\alpha_{1,c}$ parameter. If positive, then individuals in this class find the non-discolored steak more appealing in the presence of a discolored steak. If it is negative the opposite is true, and if not statistically significant then the presence of a discolored steak does not influence the appeal of the non-discolored steak. The finite mixture model is used to capture heterogeneity in the spillover effect. For example, the first class respondents might be associated with a positive spillover effect where $\alpha_{1,1} > 0$ while the second demonstrates a negative spillover effect where $\alpha_{1,2} < 0$.

The model assuming one class is estimated in four ways. First, an unweighted regression is used without demographic controls. Then, a weighted regression (without demographic controls) is used to encourage both the discolored and the non-discolored treatments (which have slight differences in their demographic composition) to display response patterns as if their demographics equaled the demographics for the entire sample. Then, the two regressions are repeated with demographic controls, including

binary variables for gender; six levels of age, income, and steak purchasing frequencies; whether they are a primary shopper; and five levels of household size. If the four models provide qualitatively different results the same four versions will be used for the model with two classes; otherwise an unweighted model without demographic controls will be used.

Statements 5 and 6 concern the willingness of a respondent to purchase the steak at a stated price, and variations in their answer in response to variations in price allow us to calculate their maximum WTP for the chosen steak by using an ordered logit model. The responses to these two statements are assumed to be generated from the random utility model specified as

(2) $$U_{i,c} = \alpha_{1,c}(D_i) - \gamma_c(P_i) + X_i\beta_c + e_{i,c}$$

where $U_{i,c}$ is the latent utility generated from being able to purchase the best steak at price P_i , D_i is an indicator variable that equals one if a discolored steak is present and zero otherwise, X_i is a vector of demographic variables, $e_{i,c}$ is an error term, and $\alpha_{1,c}$, γ_c, and β_c are parameters to be estimated. Intercepts are not needed in the model because they are implicitly reflected in the threshold parameters. An ordered logit model allows estimation of the parameters in (2). The change in the value of the non-discolored steak in the presence of the discolored steak is then calculated as

(3) $$WTP_c = \frac{\alpha_{1,c}}{-\gamma_c}.$$

If $\alpha_{1,c}$ is positive then consumers are willing to pay more for the non-discolored steak when a discolored steak is present, if negative they will pay less, and if not statistically significant the steak's value is unchanged. As before, the "c" subscript indicates that the change in the steak's value may vary across subsets of respondents.

Results

Study 1

Subjects in the internet survey were asked to look at two steaks, one discolored and one non-discolored, and were then asked to indicate whether the presence of the discolored steak makes the non-discolored steak less or more appealing by answering the two scale questions shown previously in Figure 2.1. Depending on how they answered the two scale questions, their answers might indicate (a) it makes the non-discolored steak more appealing (b) less appealing (c) or the results might be ambiguous (*e.g.*, if they indicated agreement with both statements).

The results are shown in Figure 2.2, and suggest that many more people believe they will find the non-discolored steak more appealing than those who would find it less appealing. A considerable portion of people provided an ambiguous answer, however, suggesting that they are unsure. The ambiguous responses may result from some people who answered the questions haphazardly, and some who might have been confused. They may also result from some who neither agree nor disagree that it makes the steak look more appealing but either disagree or agree that it makes the steak look less appealing. Because a number of people do not have an unambiguous opinion, and because the question asks about their conscious or stated opinions, the second study is necessary to more fully understand consumer preferences and to more fully evaluate the possible spillover effect.

Answers to the open-ended question regarding the cause of meat discoloration were analyzed by visually inspecting the answers and summarizing their content. Of the 2,596 answers, 991 correctly describe the cause of discoloration as oxidation, though of

course the words they use varied. Many actually used the word oxidation while many others referred to exposure to the air. About 340 described a process suggesting the meat is going bad and is unsafe. This includes responses indicating that bacteria causes the discoloration (123 actually used the word "bacteria"). Some correctly identified oxidation as the culprit but also mentioned that it makes the meat "go bad", and these responses are accounted for in both the 991 responses and the 340 responses. Others used the word "spoiled" or the phrase "go bad". Many respondents provided answers that were hard to categorize, like "time" or suggesting that no dye was added to the meat.

Overall these responses suggest that a large number of people do understand the cause of discoloration, and that a minority of individuals explicitly associate discoloring with food safety concerns. However, just because a person does not explicitly refer to a food safety concern does not mean that no concern is present. A person who believes oxidation is the cause of discoloring may assume that oxidation occurs because the meat was left out and exposed to the air longer than it should be, giving bacteria more opportunities to colonize the meat. Indeed, even though only 13% of people explicitly listed a food safety concern and less than that indicated the presence of bacteria, in a separate question, of all the 1,762 people who had experienced problems with meat discoloration at home, 55% say they usually throw such meat out, with 45% saying they usually eat it. This suggests that around half of the respondents do associate discoloration with either food safety concerns or simply do not eat discolored meat because of its appearance. Also, when asked whether stores should sell discolored meat, more respondents said no (49%) than those who said yes (24%), with the remainder being unsure.

This demonstrates that even though many individuals do understand the cause of discoloring, and while food scientists may not consider this cause to pose a direct safety concern, many of those individuals believe with greater exposure to oxygen also comes greater exposure to pathogens. However, since close to 60% say the discolored meat makes the non-discolored meat more attractive, they might see it as more attractive because (a) they believe the food safety threat is relegated only to the discolored meat and is not the result of a systemic management problem or (b) if they do infer a systemic management problem, the non-discolored meat is seen as a smaller safety threat.

This begs the question of what individuals mean by "more appealing". It could mean that the non-discolored steak now has a lower overall value due to a possible systemic management problem but a higher relative value compared to the discolored steak, or it could mean that the non-discolored steak actually has a higher overall value. Does "more appealing" refer to a higher overall or relative value? Study 1 cannot answer this question but Study 2 can.

Study 2

We first look at the results of the combined statements 1-4 models for the steak participants chose as their most preferred and then we turn to the results of the combined statements 5 and 6 models. As previously discussed, we estimate equation (1) first assuming homogeneous preferences (only one latent class) using ordinary least squares both with and without weighting as well as with and without the demographic control variables and compare the results in Table 2.5.

With the objective of this research in mind, the main coefficient of interest is for the *Discolor* variable. The results show that this coefficient is positive and is a relatively

small coefficient for all of the four models estimate (unweighted, weighted, unweighted with demographic controls, and weighted with demographic controls). However, it is not shown to be significant at the alpha=0.05 level for any of the models. From these results it would at first appear that the presence of a discolored steak in the choice set does not have a significant spillover effect on an individuals' preferences for the steak he/she likes most. However, it is still important to consider the possibility of unobserved heterogeneous preferences. To do this we again estimate equation (1) however we now relax the assumption of homogeneous preferences by estimating a finite mixture model (FMM) with two classes. Because the classes are unobserved (latent), there is no way to know how many classes exist. Many methods have been developed to select the number of classes to be used. However, no consensus has emerged regarding which of these methods performs best. One method which is prevalent involves minimizing the Akaike information criterion (AIC) as well the Bayesian information criterion (BIC). Generally, the FMM is estimated using multiple classes and then through comparison of the AIC and BIC between the models the one with the best fit (lowest AIC/BIC) is selected. As this method has been used extensively in the literature and is relatively simple to implement, we follow this approach to select the number of classes. Equation (1) is estimated assuming both one and two classes. Attempts to estimate the model with three classes failed to converge and therefore, we compare only the models with one and two classes. The AIC and BIC for the combined statement 1-4 model with one class (no weighting or demographic controls) are 535.46 and 543.75 respectively, while for the model with two classes they are 521.54 and 540.88. The marginally lower AIC and BIC for the two-class model favors its use over the single-class model. It is also important when comparing

FMMs to look at the significance of the coefficient for the class variable in the multinomial logit class model (used to predict probability of belonging to each class). In the two-class model the coefficient for the second class is estimated as 2.05, indicating that there is an estimated higher probability of being in the second class as compared to the first and with a p-value of 0.042 it is significant. This significant coefficient for the second class combined with the lower AIC and BIC would suggest that the two-class model is preferred to the single class for the combined statement 1-4 model. Table 2.6. displays the results from the two class FMM estimated for the combined statements 1-4.

From the results of the two class FMM it is seen that while in the first class the coefficient for *Discolor* is not found to be significant, in the second class it is significant at the alpha=.05 level. This indicates that for at least a subset of the population we would expect that the presence of a discolored steak in the choice set would have positive spillover effect on the preferences for the steak a person likes most. When we estimate the latent class marginal probability for each class we find that we would expect that 11.41% of people would belong to the first class while 88.59% of people would belong to the second class. This indicates that for a large majority of people we would expect that having a discolored steak present in their choice set would increase the appeal of the steak they have chosen as their most preferred. This finding supports the results from study 1. While the appeal of a person's preferred steak may increase when marketed together with a non-discolored steak, we still have not answered the question of whether this increased appeal translates into added value. To answer this question we will need to model the combine statements 5 and 6 and calculate the WTP premium or discount for a participant's preferred steak when the choice set includes a discolored option.

WTP Estimates

As discussed in the 'Methods and Data' section the data generating process for the combined statements 5 and 6 is best represented with an ordered logit model. The model results assuming one class (homogeneous preferences) are summarized in Table 2.7. We again relaxed the homogenous preference assumption and estimated a two class FMM, however the model failed to converge with two classes using an ordered logit model. Using least squares regression with two classes the model converges but the classes were not found to be significant (coefficient for the second class variable in the multinomial logit class model not significant) and for this reason we revert back to discussing only the results from a single class model. The results from the two class FMM (LS) can be seen in Appendix 2C.

The estimated coefficients for the price variables in both the model with and without demographic controls are negative and significant. This indicates that as the price ($/lb.) increases for the participants' preferred steak we would expect their utility to decrease. The positive sign on the coefficient for the *Discolor* variables indicates that when a discolored steak is present in the choice set a participant's utility from the preferred steak would be expected to increase. However, this coefficient is not statistically significant. Using the results from the model without demographic controls we find that consumers would be expected to pay an additional premium of $0.27, (WTP= -(0.1046/-0.3827)=0.2733), for the steak they like most when a discolored steak is present in the choice set. However, because the coefficient for *Discolor* is not statistically significant we cannot say with confidence that the actual premium would be

different from zero. This would indicate that marketing a discolored steak together with a non-discolored steak would not be expected to have a significant effect on consumers' WTP for their preferred steak.

Summary and Conclusions

Limitations

These results are not without limitations. Because actual purchase data is not used, the choices respondents make within studies 1 and 2 may contain hypothetical bias. As with any stated preference method, perhaps the largest concern comes from recognizing a lack of consistency that can exist between hypothetical behavior and behavior with real economic consequences (Penn and Hu, 2018). While there have been many methods developed to help reduce hypothetical bias within stated preference research (Blackburn, Harrison, and Rutström, 1994; Champ et al., 1997; Fox et al., 1998), it can perhaps never be fully eliminated with certainty without moving to revealed preference methods. While discoloration is not routinely sold in large quantities, additional research in this area may benefit from attempting to coordinate with retailers and collect beef purchasing data to avoid hypothetical bias and improve the validity of results.

Conclusions

The nationwide survey of US consumers used in *study 1* indicated that non-discolored steak may appeal more (positive spillover) to consumers when displayed next to discolored steaks. The in-person controlled experiment used in *study 2* supports the results of *study 1.* The results indicate that for a large subset of consumers we would expect a discolored steak present in the choice set to have a positive spillover effect on the appeal of the consumers most preferred steak. However, this positive spillover in

appeal does not seem to carry through to a positive spillover in value. When evaluating WTP we found no significant evidence that WTP for a consumer's preferred steak would change as a result of being marketed with non-discolored beef present. Thus, while the overall appeal of the preferred steak may increase when marketed next to discolored beef we do not expect consumer WTP to be effected. This would suggest that as retailers continue to attempt to minimize waste and lost revenue from discolored beef they should continue to develop internal protocols as to what to do with discolored beef, whether that be discounting the price, grinding it for hamburger, using it in prepared foods (i.e. deli products), or simply discarding it. The results of this study do not suggest that retailers should expect any negative impacts on consumer WTP from marketing discolored beef together with non-discolored beef.

As consumers have been shown to generally view discolored beef as unwholesome, more resources should be concentrated on consumer education in this area. As consumers become better educated and understand beef coloring better we could expect less waste from discoloration. Additionally, it may be of worth to push for consumer acceptance of vacuum sealed beef within the United States. Vacuum sealing greatly extends the shelf life of retail beef. However, when beef is vacuum sealed its color remains a dark purple as it is not exposed to oxygen and thus does not have a chance to bloom and take on the bright, cherry-red color. As consumers prefer the cherry-red color, additional marketing and consumer education would be needed to help consumers recognize the benefits to vacuum sealing.

Table 2.1. Questionnaire Respondent Descriptive Statistics

Variable	Definition	Percentage (all, N= 117)	Percentage (non-discolored treatment, N = 61)	Percentage (discolored treatment, N = 56)
Gender	Female	67.80	64.52	71.43
	Male	32.20	35.48	28.57
Age	18 to 21 years old	33.90	41.94	25.00
	22 to 30 years old	23.73	19.35	28.57
	31 to 40 years old	16.10	12.90	19.64
	41 to 50 years old	8.47	8.06	8.93
	51 to 60 years old	9.32	4.84	14.29
	Over 60 years of age	8.47	12.90	3.57
Income	Less than $30,000	44.44	37.70	51.79
	$30,000 to $59,000	21.37	21.31	21.43
	$60,000 to $89,999	12.82	14.75	10.71
	$90,000 to $119,999	11.11	13.11	8.93
	$120,000 to $149,999	3.24	6.56	0.00
	$150,000 or more	6.84	6.56	7.14
Shopper	Is the primary shopper in the household	68.64	59.68	78.57
	Is not the primary shopper in the household	31.36	40.32	21.43
Frequency of steak consumption	Never	14.53	13.11	16.07
	Rarely	39.32	45.90	32.14
	At least once a month	18.80	22.95	14.29
	At least once every two weeks	20.51	14.75	26.79
	At least once a week	5.98	3.28	8.93
	Almost every day	0.85	0.00	1.79
Household size	One	16.10	12.90	19.64
	Two	40.68	35.48	46.43
	Three	19.49	27.42	10.71
	Four	17.80	20.97	14.29
	Five or more	5.93	3.23	8.93

Table 2.2. Color Characteristics of Bright-red and Discolored Steak on Day of Study

Parameter	Bright-red steak	Discolored steak
CIE a* values	29.4	19.5
Percent metmyoglobin	8.2	64.5

Table 2.3. Summary of Responses to Questionnaire Statements

Statement	Response	Percent 'Most'	Percent 'Least'
1. The quality of this ribeye steak is high.	strongly disagree	0.0%	7.6%
	disagree	0.0%	18.6%
	somewhat disagree	1.7%	24.6%
	neither agree nor disagree	0.8%	9.3%
	somewhat agree	11.9%	21.2%
	agree	54.2%	16.9%
	strongly agree	31.4%	1.7%
2. I would seriously consider purchasing this steak.	strongly disagree	0.0%	15.3%
	disagree	0.0%	20.3%
	somewhat disagree	1.7%	18.6%
	neither agree nor disagree	1.7%	7.6%
	somewhat agree	11.9%	21.2%
	agree	50.0%	13.6%
	strongly agree	34.7%	3.4%
3. If offered at a reasonable price, people would purchase this steak.	strongly disagree	0.0%	5.1%
	disagree	0.0%	10.2%
	somewhat disagree	0.0%	11.0%
	neither agree nor disagree	0.9%	5.1%
	somewhat agree	2.6%	28.8%
	agree	37.9%	30.5%
	strongly agree	58.6%	9.3%
4. Most people would like the taste of this steak.	strongly disagree	0.8%	5.1%
	disagree	0.0%	11.0%
	somewhat disagree	0.8%	11.0%
	neither agree nor disagree	3.4%	21.2%
	somewhat agree	10.2%	25.4%
	agree	55.9%	22.0%
	strongly agree	28.8%	4.2%
5. and 6. I would purchase this steak for ($X)[a] per pound.	strongly disagree	3.4%	32.2%
	disagree	7.2%	22.0%
	somewhat disagree	11.4%	11.0%
	neither agree nor disagree	10.6%	9.3%
	somewhat agree	22.0%	11.9%
	agree	21.6%	9.7%
	strongly agree	23.7%	3.8%

[a] The price per pound varied randomly by $0.25 increments from $5.00 to $15.00

Table 2.4. Summary of Factor Analysis Loadings

Statement	Steak desired most Factor 1 Loadings	Steak desired least Factor 2 Loadings
1. The quality of this ribeye steak is high.	0.6477	0.7772
2. I would seriously consider purchasing this steak.	0.6373	0.8595
3. If offered at a reasonable price, people would purchase this steak.	0.6366	0.7238
4. Most people would like the taste of this steak.	0.6237	0.8217
5. I would purchase this steak for ($X)[a] per pound.	0.2557	0.6435
6. I would purchase this steak for ($X)[a] per pound.	0.3674	0.5175

[a] The price per pound varied randomly by $0.25 increments from $5.00 to $15.00

Table 2.5. Regression Results for Statements 1-4

	Unweighted Regression	Weighted Regression[a]	Unweighted Regression with demographic controls[b]	Weighted Regression[a] with demographic controls[b]
Intercept	24.573	24.451	22.761	25.457
p-value	(0.000)	(0.000)	(0.000)	(0.000)
Discolor	0.587	0.552	0.636	0.409
p-value	(0.170)	(0.257)	(0.238)	(0.449)

[a] Weights were acquired from a raking / sample balancing algorithm ensuring the weighted demographic averages for the discolored treatment and the non-discolored treatment individually matched the demographic average for the sample as a whole, for the variables gender, age, frequency of steak purchases, household size, and whether they are the primary shopper.

[b] Demographic controls included dummy variables for gender, six levels of age, income, and steak purchasing frequencies; primary shopper, and five levels of household size.

Table 2.6. Results of Combined Statements 1-4 Two class Finite Mixture Model

	Class 1		Class 2	
	Coefficient	P-value	Coefficient	P-value
Intercept	21.0972	0.000	24.9414	0.000
Discolor	0.3939	0.902	0.7881	0.038
Probability of Respondent Belonging to Class	11.41%		88.59%	

AIC:521.54 BIC:540.88

Table 2.7. Ordered Logit Results for Statements 5 and 6

	Unweighted Model	Unweighted Model with Demographic Controls[b]
Discolor	0.105	0.082
	(0.657)	(0.768)
Price	-0.383	-0.408
	(0.000)	(0.000)

[a] Threshold parameters are not shown.
[b] Demographic controls included dummy variables for gender, six levels of age, income, and steak purchasing frequencies; primary shopper, and five levels of household size.

Discolored Steak Regular Steak

Above is a discolored steak to the left and a regular steak to the right. Imagine you saw these two steaks side-by-side in a grocery store.

	Strongly agree	Agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Disagree	Strongly disagree
The discolored steak on the left makes the regular steak look less appealing than it really is.	○	○	○	○	○	○	○
The discolored steak on the left makes the regular steak look even more appealing than it really is.	○	○	○	○	○	○	○

Figure 2.1. Grid Style Question from Nationwide Consumer Survey

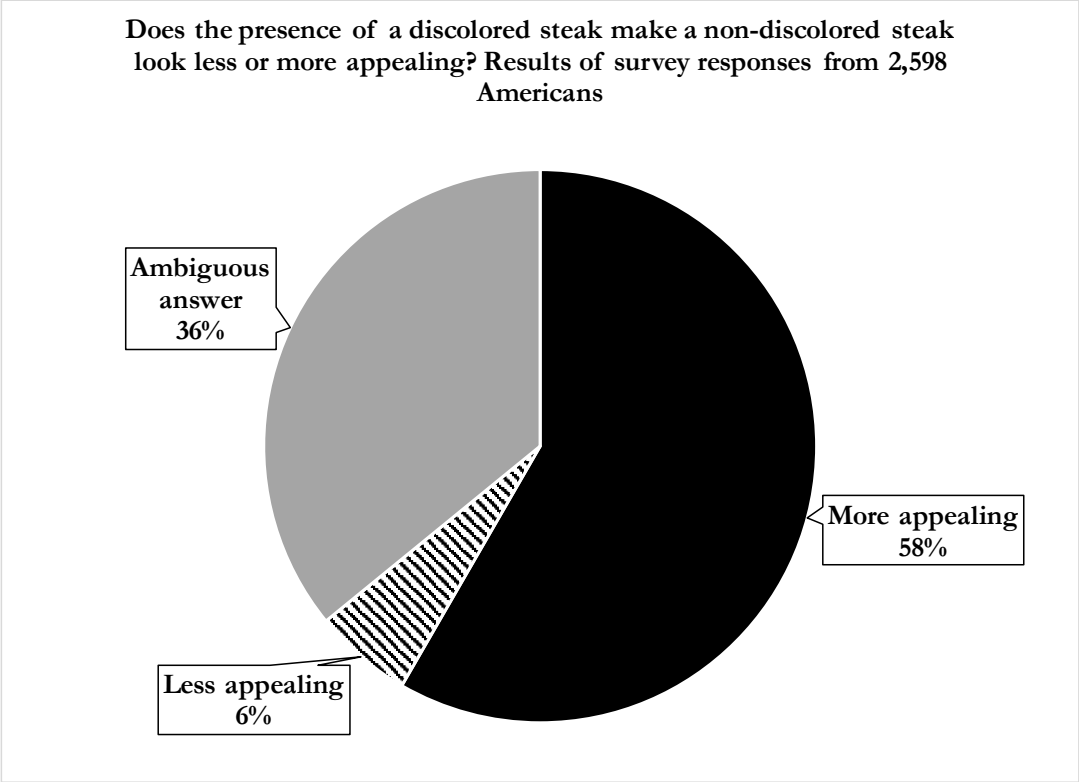


Figure 2.2. Summary Results of Survey Question Asking About Appeal of Steak

CHAPTER III

INTRODUCING DESIGN VALUATION WITH APPLICATION TO VALUING COLLEGE GRADUATE ATTRIBUTES

Abstract

Design valuation is a new valuation method where subjects design their optimal good by selecting various attributes at select prices. Through a design valuation survey of college graduate employers, willingness-to-pay (WTP) data are collected for various college graduate attributes. Average WTP estimates for the college graduate attributes are estimated relative to the type of college (agricultural, business, engineering, or other) from which the employer prefers to hire recent graduates. The estimates from the various college types are then compared and contrasted, aiding both academic advisors and students in adjusting their current expectations and preparing for the job market.

Key Words: interval-censored data, non-market valuation, academic advising, human resources

Introduction

The total value of a good is often assumed to be the sum of the values of its components. Following this logic, we could similarly describe the value of a recent college graduate as a sum of the individual values of the attributes the graduate possesses. As employers seek to hire recent graduates they may value some attributes more than others. Additionally, the values employers place on these attributes is likely heterogeneous and could differ significantly based on individual employer characteristics.

There has been a significant amount of research evaluating the relative importance of various college graduate attributes and skills towards employability. Suleman (2016) compared several findings of past research and demonstrated that while wide agreement exists on the need for relational skills, namely interpersonal, communication and teamwork abilities, the evidence showed little consensus on which skills best foster employability. Using a web-based choice experiment surveying California area agribusiness employers Noel and Qenani (2013) highlighted a trend in the needs for skills of agribusiness graduates, with skills such as creativity and critical thinking becoming quite important in the labor market. In addition to studies measuring the relative importance of graduate attributes there have also been numerous studies which have estimated the value for recent college graduate attributes using various techniques (Barkley 1992; Barkley et al. 1999; Norwood and Henneberry 2006). The past literature has approached this topic using different methods. Barkley (1992) and Barkley et al. (1999) used survey data collecting salary information of recent graduates and regressed this on individual attributes to estimate the value of specific attributes. Norwood and Henneberry (2006) used a choice experiment to value recent graduate

attributes by presenting respondents with job candidates who had differing attributes and salaries. The purpose of this study is to present a new method of stated preference elicitation called design valuation as well as to estimate the value employers place on various college graduate attributes. We add to the literature on recent college graduate attribute valuation by further classifying attribute values by employer type. Four employer type classifications are used; those that prefer to hire graduates from an agricultural, business, engineering, or other college. The value estimates for specific attributes can then be compared across the differing employer types and we highlight differences and similarities.

A wealth of economic studies have developed, tested, and refined tools for measuring stated preferences through surveys (see Lusk and Hudson 2004 for an overview). Two methods are predominately used in this literature – conjoint analysis and contingent valuation. The goods evaluated with these survey instruments are defined as a collection of attributes. For example, a lake may be described by water clarity, frequency of algae blooms, and boat ramp access; and steaks may be described by their tenderness, marbling, and days of carcass aging.

In using these stated preference methods, the researcher designs various goods by assigning each good a unique collection of attributes. Subjects are then asked to indicate their preferences for these goods. Subjects are involved in the research in a post-design stage, after the good has been designed. *Post-design valuation* is often touted because it mimics many real decisions, such as which brand of flour to purchase and whether to approve a referendum providing a public good.

However, consumers often face real decisions in the pre-design phase, decisions which can be mimicked using a process we call *design valuation*. Households purchasing a new home will often design it themselves by selecting the attributes they prefer; such as ceiling height, number of bathrooms, and number of stories. The particular attributes chosen are determined by both preferences and attribute prices. Consider an alternative example, a computer upgrade. Assume a marketing researcher is interested in the values a consumer places on different computer components or upgrades (e.g. larger monitor, more powerful processor, etc.). A conjoint approach would present the consumer differing upgrades above the baseline computer at varying prices. The upgrade is a collection of attributes with unique prices for each attribute collection. This can pose a large cognitive burden if the attribute list is long and the number of alternatives to peruse large. Imagine having to keep track of five different upgrades each described by a unique combination of 15 attributes. An alternative approach is to allow consumers to purchase individual attributes at varying prices. The consumer can then consider each attribute at a time, compare it to the attribute price, and decide whether to include this attribute in their virtual shopping cart.

This is the fundamental idea behind the design valuation method used in this study. Design valuation operates by defining a general good as a collection of attributes and assigning prices to those attributes. Respondents are then asked to design their optimal good based on those attribute prices (much like a customer designs their optimal personal computer). By varying the attribute prices within and across surveys, the value of each attribute can be inferred.

In a sense, design valuation is similar to asking multiple contingent valuation questions. Returning to the computer example, the marketing researcher may ask the customer if she would purchase each individual upgrade at the stated price, which is analogous to one contingent valuation question per upgrade component. Each purchase would be in addition to the baseline computer at a base cost. If these multiple contingent valuation questions were provided in a mail or phone survey, this process may be too difficult for the customer. It is possible that the customer would not be able to easily track her total price of the computer, what upgrade she had previously purchased and the price of the upgrade, and changing previous selections may not be possible. Designing her optimal computer would be problematic.

To alleviate this problem, our proposed design valuation survey is internet-based where a built-in calculator relieves the subject of this burden, providing the individual with a more direct and concise question. Design valuation has no obvious statistical advantages over post-design methods. If humans were perfectly rational, had well-defined preferences, did not suffer from survey fatigue, and had perfect memory, both design and post-design methods would elicit identical preferences. However, design valuation is preferred over traditional post-design valuation methods in this paper because of its practicality; its ability to extract much information from a simple question.

To achieve the purpose of this study an internet-based design valuation survey is constructed and sent to employers of college graduates. The attributes evaluated resemble those in Norwood and Henneberry (2006), Boland and Akridge (2004), Berle (2007), and Litzenberg and Schneider (1987).

Norwood and Henneberry (2006) employed a choice based conjoint survey or a *post-design* survey to estimate employer's willingness-to-pay for college graduate attributes. Therefore, their results can be compared to the results from our design valuation method. The specific design valuation format used provides interval-censored willingness-to-pay data (an interval known to contain the individual's true value) on college graduate attribute values. Using the interval-censored data collected we use interval regression to estimate the value that employers place on the specific college graduate attributes.

Theory

Any good can be thought of as a collection of attributes and the goods' value a function of the individual attribute values (Rosen 1974). Let a hypothetical good be described by attributes a_i $(i = 1, \dots, I)$. An attribute a_i may be a dummy variable indicating the presence or absence of some trait (e.g. excellent communication skills), or a continuous variable denoting the level of some attribute (e.g. grade point average). Only the dummy variable case is considered here. Further, let the value of attribute a_i to an individual be denoted v_i, assumed independent of other attributes, and stated in money metric form. The value of good *j* is then measured by the function $\sum_{i=1}^{5} a_{i,j} v_i$ where a_{ij} refers to the presence or absence of attribute *i* in good *j*. If the price of the good *j* is P_j, the surplus received from good *j* is

(1) $$U_j = \sum_{i=1}^{5} a_{ij} v_i - P_j.$$

Assuming a consumer of good *j* is a welfare maximizer the optimization problem the consumer faces is

(2) $$\max_{a_{ij}} U_j = \sum_{i=1}^{5} a_{ij} v_i - P_j.$$

Post-design valuation methods such as contingent valuation and conjoint analysis utilize questionnaires to determine if U_1 is less than, equal to, or greater than U_2. Researchers only observe the sign of $U_1 - U_2$, and from these observations must infer the values of the v_i's. For example, suppose a respondent is asked to choose one of the following two goods: good 1 ($a_{11} = 1,\ a_{21} = 1, a_{31} = 1, a_{41} = 0, a_{51} = 0, P_1 = 1$) or good 2 ($a_{12} = 0,\ a_{22} = 0, a_{32} = 0, a_{42} = 0, a_{52} = 0, P_2 = 0$). This particular choice resembles a contingent valuation question where the respondent is asked if she would like a public good provided that it increases taxes by 1. If good 1 is chosen, all the researcher knows is that $v_1 + v_2 + v_3 \geq 1$.

Now consider a design valuation question where the individual is given the baseline good or good 1 ($a_1 = 0,\ a_2 = 0, a_3 = 0, a_4 = 0, a_5 = 0, P = 0$) and is allowed to purchase any attribute a_i at a price of 0.4. Suppose attributes a_1 and a_2 are purchased, revealing to the researcher that $v_1 > 0.4$, $v_2 > 0.4$, and $v_3, v_4,$ and $v_5 < 0.4$. Clearly more information is obtained in the design valuation question. This does not imply that design valuation is necessarily superior to *post-design* methods. The same information could be obtained through five contingent valuation questions; one comparing ($a_{11} = 1,\ a_{21} = 0, a_{31} = 0, a_{41} = 0, a_{51} = 0, P_1 = 0.4$) to ($a_{12} = 0,\ a_{22} = 0, a_{32} = 0, a_{42} = 0, a_{52} = 0, P_2 = 0$), another comparing ($a_{11} = 0,\ a_{21} = 1, a_{31} = 0, a_{41} = 0, a_{51} = 0, P_1 = 0.4$) to ($a_{12} = 0,\ a_{22} = 0, a_{32} = 0, a_{42} = 0, a_{52} = 0, P_2 = 0$), and so

on. In fact, design valuation can be thought of as a series of contingent valuation questions, where the contingent valuation question is posed for each attribute and the respondent makes their design for each attribute jointly in the same general question.

The statistical information gleaned from a design valuation question will then be equivalent to a number of contingent valuation questions. The advantage of design valuation is that it contains those contingent valuation questions in one compact question, easily answered in internet browsers. The design valuation format will also be familiar to consumers who often design their own product through a manufacturer whether it is a new car, a computer, or a home. Subjects should be able to perform the design valuation task with little instruction, ensuring high response rates and greater information.

Data

In the fall of 2006 employers of Oklahoma State University graduates were asked to participate in an internet survey eliciting their preferences for new hires. The invitations were mailed to 4,401 employers yielding 507 responses for a response rate of 12 percent. This is similar to the response rate of employers conducted by Norwood and Henneberry (2006). Unlike the Norwood and Henneberry study, we do not restrict the list of employers to just those who are known to hire agricultural graduates, but employers of all undergraduate degrees. This will allow for not only the estimates for WTP to be made for those hiring graduates from agricultural colleges, but other colleges as well which may provide insightful information as comparisons are made.

Refer to Figure 3.1 illustrating one of the design valuation questions used in this study. Employers of college graduates are presented with a baseline graduate requiring a $25,000 salary, and possessing low levels of five attributes. The employer is then

allowed to purchase any of the five attributes at different prices. The cognitive burden of this question is relatively low, especially considering it is the equivalent of five contingent valuation questions.

Given that *post-design* valuation tools have been used for decades, it is likely that valuation practitioners have considered using design valuation. However, given that most valuation studies use mail or phone surveys, the fact that respondents would have to calculate the total good cost based off their attributes purchased likely made design valuation seem impractical. Internet surveys are the ideal platform for hosting design valuation questions because automatic calculators can be easily installed in the survey software. See Figure 3.1 where the respondent can click one button and have the computer recalculate the salary based off the attribute purchases.

Contingent valuation questions often have a similar follow-up question, where if the good is purchased (not purchased) in the first question its price is raised (lowered) in the second. This is referred to as dynamic updating. The questions are dynamic in the sense that one question depends on the answer to a previous question. Following the previous example, since good 1 $(a_{11} = 1,\ a_{21} = 1, a_{31} = 1, a_{41} = 0, a_{51} = 0, P_1 = 1)$ is preferred to good 2 $(a_{12} = 0,\ a_{22} = 0, a_{32} = 0, a_{42} = 0, a_{52} = 0, P_2 = 0)$, the respondent can be asked to make the same choice where P_1 is increased to 2. Since the value of multiple attributes is of concern, researchers would rarely repeat the same combination of attributes as in this example. And while algorithms are available for developing dynamic attribute combinations in internet surveys, they are complex and difficult to program in traditional browser software.

Dynamic updating is straightforward in design valuation though. Refer again to Figure 3.1. Suppose that the respondent chooses to purchase number crunching skills at $500 and ability to work well with others at $17,500, but none of the other attributes. A follow-up question would then increase the price of number crunching skills and ability to work well with others while lowering the price of the remaining attributes. Such dynamic updating is employed in the survey, producing data on attribute values that are interval-censored. For example, if the respondent purchased number crunching skills for $500 in the first question but declined in the second question when the cost rose to $5,000, the interval-censored observation is ($500, $5,000). The true value of this attribute for the employer is known to reside within this interval. If it is purchased at both prices, the interval is $5,000 and an upper bound. If an attribute is not purchased at either price, assuming attribute values are non-negative, the interval is zero and the lowest price offered of $500.

The first page of the survey informed the employer that the purpose of the survey was to seek input on what kind of college graduate they prefer to hire, and asked them to answer questions in a manner that best reflected their actual hiring practices. On the second page, a simple practice question was presented to help prepare the employer respondent for the more complex design valuation (DV) questions later in the survey. Before the employer was asked to answer questions similar to Figure 3.1, an information script was provided on the third page. This script provided information on the DV questions and how to answer them. For example, they were told to assume that the graduate holds a degree from a four-year educational institution, and to assume the graduate's possession of an attribute that is not listed is at an "average" level. They were

also told that if they would not hire any college graduate at the $25,000 salary level to leave the questions unanswered. Since it is impossible to distinguish these employers from those who simply did not wish to answer the questions, those responses are not included in the data analysis.

The fourth page contained the first DV question for Attribute Set A; which included internship or work experience (as opposed to no internship or work experience), one high quality academic award (as opposed to none), ability to speak and write in numerous languages, including Spanish (as opposed to only English speaking and writing skills), one high leadership position in an academic organization (as opposed to none), and outstanding letters of recommendation (as opposed to mediocre letters). Employers were first presented with a low-quality graduate earning a $25,000 salary with none of the aforementioned attributes. They were then allowed to purchase each attribute at a particular price.

Some employers have direct control over the salary they offer. Others, such as government agencies, have a set salary they must pay and hire the most qualified applicant they can obtain at this salary. For these latter employers, they will select the attributes that they deem both most important and affordable, up to their preset salary. This resembles someone who can spend no more than $1,000 on a computer upgrade, and purchases the computer part upgrades they value most and are affordable until any additional computer part upgrade exceeds the $1,000 limit.

After making their attribute selections, the fifth page presented a similar question where only the attribute prices differed. If an attribute was purchased in the previous question, its price was increased by a randomly selected percentage on the 1-100 percent

interval. Otherwise, its price was decreased by the same random percentage. The purchase decision for any one attribute on the two design valuation questions (and information on the lower and upper bound of WTP) provides an interval known to contain the employers' true WTP value.

A second dynamic element was also introduced in the survey. Internet surveys have the advantage that the survey can be modified continually as it is being taken. The initial price of each attribute was increased or decreased for each successive respondent based off respondents' willingness to purchase the attribute in prior surveys. For example, if more than 50 percent of respondents purchased internship experience, its initial price was increased on subsequent surveys. The initial price would increase across surveys until less than 50 percent purchased it, at which point the initial price would begin to decline. While the survey was administered the initial price would drift up and down such that on average 50 percent purchased the attribute, increasing the statistical efficiency of the survey design. The degree to which attribute values increased or decreased varied across attributes. Attributes whose values were hypothesized to be lower increased / decreased in $200 increments while others used increments of $500. Hypotheses of attribute values were based mainly on the Norwood and Henneberry (2006) study.

Pages 7 and 8 contain two similar DV questions designed to elicit the value for Attribute Set B shown in Figure 3.1. These include number crunching ability, character, ability to work well with others, communication skills, and problem solving abilities. The remaining questions regarded other information, such as the type of employer, size of employer, and the type of college degree they prefer to hire. Also included was a

question of whether they had influence over hiring decisions. If they did not, their responses were not included in the analysis. Excluding these and those who did not purchase any attributes reduced the sample size from 507 to 453.

Summary statistics on the employers are provided in Table 3.1. Most employers identified themselves as a government organization, a manufacturer, or other. Almost half are large employers with over 500 full-time employees. Employers were presented with a list of degrees and were asked to select their one preferred degree. These degrees were accounting; business; communications; finance; economics; management; marketing; agricultural engineering; agricultural communications; agricultural economics / agribusiness; agronomy; animal science; food science; horticulture; civil, electrical, mechanical or chemical engineering; industrial engineering; other. From these preferred degrees, we then grouped employers into four categories according to the type of college (agricultural, business, engineering, or other) from which they prefer to hire graduates.

Procedure

Employer responses to the design valuation questions were used to construct interval-censored willingness-to-pay (ICWTP) data for each attribute and employer. For example, employer *i's* value for a particular attribute is given by the interval (L_i, U_i) where L_i and U_i are the lower and upper bounds of the attribute value, respectively. Recall that each employer was given the opportunity to purchase each attribute at two different prices. For employers who purchased an attribute at one price but not another, the values of L_i and U_i are taken directly from those two prices. For employers who declined the purchase at both prices, it is assumed that $L_i = 0$ and U_i equals the lowest of those two prices. Finally, for those who purchased the attribute at both prices, L_i equals

the larger of the two prices, and U_i is set equal to the largest value of L_i for other employers. Thus we would expect that for each attribute j the true value of the attribute of each employer i resides within the constructed interval $[L_{ji}, U_{ji}]$ but is unobservable or latent. To estimate this latent value for each attribute we could simply use the midpoint of each interval, however, as noted by Stewart (1983) this method would generally result in inconsistent estimates. Stewart (1983) outlined different approaches to yield maximum likelihood estimates under the assumption of normality. STATA's "intreg" command facilitates MLE of interval response data in the case of normally distributed errors.

Thus, the data generating process for this study is

$$V_{ji}^{*} = \beta_{0j} + \beta_{1j} Buscollege_i + \beta_{2j} Engcollege_i + \beta_{3j} Othercollege_i + \varepsilon_{ji} \quad (3)$$

where V_{ji}^{*}, is the true (latent) average value of attribute j in a recent college graduate when being hired by the ith employer, $Buscollege_i$ is a dummy variable equal to 1 when the ith employer most often prefers to hire graduates from a business college and equal to 0 otherwise, $Engcollege_i$ is a dummy variable equal to 1 when the ith employer most often prefers to hire graduates from an engineering college and equal to 0 otherwise, $Othercollege_i$ is a dummy variable equal to 1 when the ith employer most often prefers to hire graduates from a college other than agriculture, business, or engineering college and equal to 0 otherwise, and with $\varepsilon_{ji} \sim N(0, \sigma_{ji}^2)$. To avoid the dummy variable trap, no variable is included for when the ith employer most often prefers to hire graduates from an agricultural college. Thus, the constant β_0 can be interpreted as the value for attribute j when the ith employer most often prefers to hire graduates from an agricultural college. The estimates for $\beta_{2j}, \beta_{3j,}$ and β_{4j} can be interpreted as the value premiums or discounts

associated with attribute j when the ith employer most often prefers to hire graduates from a business college, engineering college, or other type of college respectively.

Results

The average value of each college graduate attribute was estimated using MLE as outlined in equation (3). The value estimates for each attribute along with their estimated standard errors for employers who most often prefer to hire graduates from agricultural colleges are summarized in Table 3.2. The table also contains the attribute value premiums or discounts estimated for employers who typically prefer to hire from other colleges.

Consider the types of attributes valued. The first five (internship experience, one high quality award, foreign language, held leadership position, and recommendation) are all tangible attributes in the sense that they are easily verifiable and measurable. The second set of five attributes (number crunching ability, high degree of character, work well with others, excellent communication, and problem solving ability) are intangible in the sense that they are unmeasurable and require the employer's subjective judgement to evaluate. A quick review of the results table and it is quite apparent that on average the intangible attributes have a much higher value to employers than the tangible attributes. The larger average values as well as greater variability within the intangible skills or attributes as compared to the tangible attributes is not unexpected. Velasco (2012) demonstrated that intangible attributes (soft skills) are the most required attributes in the hiring process. Additionally, we expect that intangible skills may have different interpretations for each employer and as such would be subject to greater heterogeneity than the more tangible attributes which require much less subjectivity. As shown by

Briggeman et al. (2007), the assessment of these intangible attributes is most critically accomplished through a personal interview by the potential employer.

One academic award has the lowest value ($381.32) for any attribute. The ability to work well with others was valued the highest ($17,920) with a high degree of character and excellent communication also valued quite highly at ($17,366 and $17,464 respectively). It is of interest to note statistical differences between estimates for employers who prefer to hire from agricultural colleges as compared to other colleges (business, engineering, other). These comparisons are useful for advisors in that they better equip them with the knowledge they need to help students develop specific attributes that will result in students achieving the greatest payoff or return on their investment in education.

In comparing the results by what type of college employers prefer to hire from, we see that for many of the attributes there are no statistical differences (at 0.05 significance level). This simply indicates that we do not have enough evidence to suggest that the value estimated for employers who prefer to hire from agricultural colleges (the omitted dummy variable) would be different than the value for employers who prefer to hire from the other college types. Looking first at the tangible attributes the only statistical differences we see are for the internship experience attribute as well as the held leadership position attribute. Based on our estimates we expect employers that prefer to hire from engineering colleges to place a higher value ($5,383.47 premium) on the internship experience attribute as opposed to employers that prefer to hire from agricultural colleges. This would result in the total expected value for the internship experience attribute to be $21,064.90 for employers who prefer to hire from engineering

colleges. This indicates that relevant past experience would be expected to garner a larger premium within engineering careers. A similar comparison can be made for the leadership position attribute, only in this case we would expect a discount of $1,160.63 for employers who prefer to hire from other colleges as compared to agricultural colleges. Within the intangible attributes we see a greater variability between the different types of employers. Significant differences are found within all but the 'work well with others' attribute. The 'problem solving ability' attribute particularly shows a significant amount of heterogeneity among the employer classification types. The estimate value for this attribute for employers who prefer to hire from agricultural colleges is $14,638. This value is significantly less than the value estimated for the other three employer classifications. This indicates that as opposed to employers who prefer to hire from business, engineering, or other colleges we would expect employers who prefer to hire from agricultural colleges to place less value on problem solving ability in recent graduates they are looking to hire. This may suggest that if only looking to maximize their return on their investment in education students who possess a high level problem solving ability may be better served by not majoring within agriculture.

While the average value of foreign language skills is $1,376, this is not the best rate-of-return estimate for students acquiring this skill. The average value refers to all employers, those who do and those who do not need employees that know Spanish and other languages. Those acquiring a language will be more likely to interview at jobs where multiple language skills are stressed, which is consequently those employers who place higher values on this skill. For students considering learning Spanish, a rate-of-return higher than the average would be expected. The same argument can be made for

number crunching abilities. Many jobs do not require employees to possess quantitative skills, thus the relatively low average value of $2,473. Yet, those who acquire this skill will likely interview with employers who place a greater than average value on this skill, and thus could expect to receive a greater value for their number crunching ability than the average value. To get a better idea of the distribution for these attributes we use the Turnbull estimator which is best described as a nonparametric maximum likelihood estimator (Turnbull 1974).

Suppose that observation i contains a lower bound L_i and an upper bound U_i known to contain the true value willingness-to-pay (WTPi). The Turnbull estimator first requires one to order the L_i and U_i (stacked in the same column) values in ascending order. Then, identify those intervals (L_i , U_j) (where j can equal i but does not have to) for which no other lower or upper bound are captured within this interval. These are called equivalence classes, and are the only intervals over which the likelihood can assign probability mass (Day 2007).

Suppose these equivalence classes are denoted $C_0 < C_1 < C_2 <, \dots, < C_E$. The Turnbull estimator operates by estimating the cumulative distribution for WTP_i. Specifically, the cumulative distribution at each C_i, denoted $F(C_i)$, is estimated by maximizing the log-likelihood function

(4) $$LLF = \sum_{i=1}^{I} \ln\left[\sum_{e=1}^{E} d_{ie}\left(F(C_e) - F(C_{e-1})\right)\right]$$

where $d_{ie} = 1$ indicates the WTP interval (L_i , U_i) spans the equivalence class (C_{e-1} , C_e). The optimization routine must be constrained so that $0 < F(C_0) < F(C_1) <$, … , $< F(C_E) < 1$.

After estimation of the CDF for both language skills and number crunching ability we see that, as expected, for the majority of employers we would estimate their value for these attributes to be much lower than the mean. However, the CDF estimates for both of these attributes also indicate that for a minority of employers their value for these attributes is much greater than the average. For language skills the estimated CDF would indicate that nearly 49 percent of employers would be expected to value language skills at less than $255. However, we would expect that approximately 15 percent of employers would value language skills at above $5,000 and 10 percent of employers would value this skill at over $8,000. This same pattern appears within the estimated CDF for the number crunching ability attribute. Approximately 50 percent of employers would be expected to value number crunching ability at less than $540. However, graduates with this skill need not despair as 25 percent of employers indicate that they value this skill at over $6,400 and 13 percent of employers value it over $10,000. Thus, for these attributes where the dispersion of employer values is quite large we would expect those graduates with these attributes to seek out employers who value what they have to offer and would likely realize much greater returns than the average value estimates would otherwise indicate.

As the study conducted by Norwood and Henneberry (2006) employed a choice based conjoint survey or a *post-design* survey to estimate employer's willingness-to-pay for college graduate attributes, their results can be compared to the results from our design valuation method. The values for internship experience, character, and communication skills are consistent with those values calculated by Norwood and Henneberry (2006) who use a traditional choice experiment and conventional estimation

techniques. The average values reported in Table 3.2 are roughly two times larger than those reported in Table 5 of Norwood and Henneberry (2006, pg. 492). However, the value in Norwood and Henneberry (2006) assume there are multiple job candidates to choose among, and the values in Table 3.2 assume only one candidate that can be designed to the employer's desires—the more conventional calculation. Using the parameter estimates in Table 4 of Norwood and Henneberry (2006, p. 490), and the more conventional value calculation, their value estimates for internship experience, character, and communication skills are $22,000, $39,430, and $35,602, respectively, which are either equal to or greater than the values reported in Table 3.2. Moreover, Norwood and Henneberry report average values for one academic award and one leadership position of $663 and $2,406, both of which are similar to those reported in Table 3.2.

Summary and Conclusions

Design valuation is a unique survey method that allows respondents to participate in the *pre-design* survey process. Respondents are given a general good, described by various attributes, and are allowed to change the attribute levels at prescribed prices. The respondent, therefore, designs the good herself. Design valuation is equivalent to a number of contingent valuation questions; the two are statistically equivalent but are implemented differently. Thus, the preference of one design valuation or multiple contingent valuation questions depends on the practicality and the cognitive burden posed on the respondent.

Future research should measure the cognitive burden of each approach and respondent preferences for the two methods. Similar research could be expanded to compare design valuation to conjoint analysis. Essentially, we suggest that researchers

measure preferences for stated preference instruments, by those responding to the instruments. If two methods elicit the same degree of information but one is answered more easily by the respondent, that instrument should receive some preference.

Future studies should also measure the extent to which design valuation is subject to anchoring. It is well known that in double-bounded contingent valuation questions, individual values depend on the initial prices posed. Such biases would then be expected in design valuation as well. Yet, even single-bounded contingent valuation is subject to anchoring, so conjoint analysis may be as well. More information on the presence of anchoring under these three alternative formats is desirable.

Finally, the attribute value estimates beg a number of questions. Items like ability to work well with others is valued highly, but what exactly does this mean? Does it imply ability to engage in stimulating conversations, general manners, or emotional intelligence as being increasingly studied by psychologists, etc.? Similarly, although problem-solving abilities are highly valued, what type of problems are employers thinking of when they complete the design valuation questionnaire? Finally, when employers indicate they value "high character", what percent of college graduates do they perceive have such high character? For academic advising to progress from the use of the values estimated in this paper, these questions warrant attention.

Despite these limitations, the current study provides useful information for advisors and students alike. Of most importance would be the ability to compare and contrast the importance (magnitude of value estimates) of the college graduate attributes for the employers who prefer to hire from specific types of colleges. Results indicate the highest attribute values from employers who prefer to hire from agricultural colleges are

the ability to work well with others, excellent communication skills, and a high degree of character. Additionally, problem-solving ability is valued the least among employers who prefer to hire from agricultural colleges as compared to the other colleges, indicating that graduates with this attribute would be better suited for majors outside of agriculture. Conversely, results indicate that number crunching ability is valued highest among employers who prefer to hire from agricultural colleges indicating agricultural college students can expect a better than average return for this attribute as compared to their nonagricultural college student peers. This information, along with the other values for recent college graduate attributes estimated in this study, allow students to better align their own goals in the development of specific skills and attributes to increase their marketability and return on education investment upon entering the job market. This information also benefits college advisors. The comparison of attribute value estimates by employer classification type better equips advisors to guide students to pursue specific degrees based off of the qualities they possess. As students graduate and seek employment they must market themselves according to the talents, skills, past experience, and abilities they possess to be successful. This research aids in adjusting students' expectations for what attributes employers truly value in hiring recent graduates.

Table 3.1. Employer Demographics (Sample Size = 453)

Organization Type	Percent	Preferred Degree	Percent Preferred Degree to Hire From	Number of Full-Time Employees	Percent
Government Organization	15%	Accounting	6%	< 10	4%
Manufacturer	20%	Business Communications	4%	10-49	16%
Financial Service Provider	9%	Finance	4%	50-59	13%
Consultant	10%	Economics	0%	100-500	22%
Food Processor	2%	Management	8%	> 500	45%
Retailer	4%	Marketing	6%		
Wholesaler	3%	Ag Engineering	2%		
Farm or Livestock Producer	2%	Ag Communications	1%		
Farm Input Supplier	3%	Ag Econ / Ag Bus	8%		
Other	32%	Agronomy	3%		
		Animal Science	4%		
		Food Science	1%		
		Horticulture	1%		
		Civil, Electrical, Mechanical, or Chemical Engineering	22%		
		Industrial Engineering	3%		
		Other	24%		

Notes: numbers may not sum to one due to rounding.

Table 3.2. Value for Recent College Graduate Attributes with respect to the Type of College the Employer Prefers to Hire From

	Value If Hired From an	Change in Value Based on Employer Category		
Attribute	Agricultural College	Business College	Engineering College	Other College
Internship experience	$15,681	$2,233	$5,383*	-$689
	(1741)	(2158)	(2251)	(2241)
One high quality award	$381	-$100	$49	$77
	(66)	(82)	(86)	(87)
Foreign language	$1,376	-$191	-$279	-$379
	(212)	(260)	(268)	(272)
Held leadership position	$2890	-$467	-$376	-$1,161*
	(288)	(356)	(371)	(369)
Recommendation	$2,392	-$406	$40	-$36
	(294)	(360)	(379)	(381)
Number crunching ability	$2,473	-$491	-$262	-$1,118*
	(302)	(369)	(390)	(383)
High degree of character	$17,366	$4,542	$6,917*	$6,372*
	(2319)	(2874)	(2991)	(3022)
Work well with others	$17,921	-$1,767	$1,986	-$161
	(1854)	(2295)	(2405)	(2417)
Excellent communication	$17,464	$4,033	$8,744.87*	$4,512.86
	(2329)	(2878)	(3026.11)	(3021.48)
Problem solving ability	$14,638	$6,527*	$11,733*	$8,551*
	(2274)	(2817)	(2953)	(2970)

Notes: Numbers reported in parenthesis are standard errors.
* indicates estimates that are significantly different at 0.05 level from value if employer prefers to hire from agricultural college

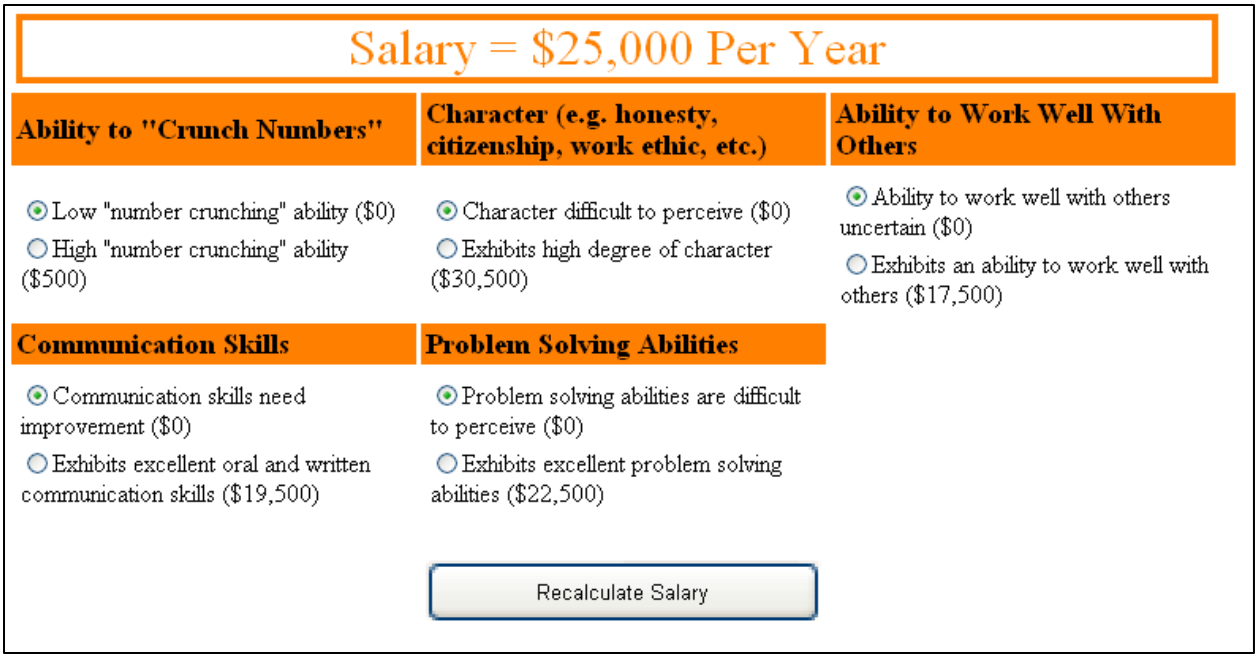


Figure 3.1. Design Valuation Question

REFERENCES

Barkley, A. P., W. A. Stock and C.K. Sylvius. 1999. Agricultural Graduate Earnings: The Impacts of College, Career, and Gender. *American Journal of Agricultural Economics* 81:785-800.

Barkley, A.P. 1992. Earnings of Kansas State University Agriculture Graduates: 1978-88. *American Journal of Agricultural Economics* 74(1):215-222.

Berinsky, A.J., M.F. Margolis, and M.W. Sances. 2014. Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys. *American Journal of Political Science* 58: 739-753.

Berle, D. 2007. Employer Preferences in Landscape Horticulture Graduates: Implications for College Programs. *NACTA Journal* 51(1):21-25.

Beshears, J., J.J. Choi, D. Laibson, B. Madrian, and B.C. Madrian. 2008. How are preferences revealed? *Journal of Public Economics* 92: 1787-1794.

Blackburn, M., G.W. Harrison, and E.E. Rutström. 1994. Statistical Bias Functions and Informative Hypothetical Surveys. *American Journal of Agricultural Economics* 76(5): 1084-1088.

Boland, M.A. and J.T. Akridge. 2004. Undergraduate Agribusiness Programs: Focus or Falter? *Review of Agricultural Economics* 26(4): 564-578.

Briggeman, B., S.R. Henneberry, and F.B. Norwood. 2007. How do employers assess job candidate attributes? *NACTA Journal* 51(3): 15-21.

Cannell, C.F., and R.L. Kahn, 1968. Interviewing, in Lindzey, G., and E. Aronson (eds), Handbook of Social Psychology, Vol. 2. Reading, MA: Addison Wesley.

Carpenter, C.E., D.P. Cornforth, and D. Whittier. 2001. Consumer Preferences for Beef Color and Packaging did not Affect Eating Satisfaction. *Meat Science* 57(4): 359-363.

Champ, P.A., R.C. Bishop, T.C. Brown, and D.W. McCollum. 1997. Using Donation Mechanisms to Value Nonuse Benefits from Public Goods. *Journal of Environmental Economics and Management* 33(2): 151-162.

Cummings, R.G., and L.O. Taylor. 1999. Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *American Economic Review* 89: 649-665.

Day, B. 2007. Distribution-Free Estimation with Interval-Censored Contingent Valuation Data: Troubles with Turnbull? *Environmental and Resource Economics* 37(4):777-795.

Demartini, E., D. Vecchiato, T. Tempesta, A. Gaviglio, and R. Viganò. 2018. Consumer Preferences for Red Deer Meat: A Discrete Choice Analysis Considering Attitudes Towards Wild Game Meat and Hunting. *Meat Science* 146: 168-179.

Faustman, C., and R.G. Cassens. 1989. Strategies for Improving Fresh Meat Color. *International Congress Meat Science and Technology* 35:446-453.

Feuz, D.M., W.J. Umberger, C.R. Calkins, and B. Sitz. 2004. U.S. Consumers' Willingness to Pay for Flavor and Tenderness in Steaks as Determined with an Experimental Auction. *Journal of Agricultural and Resource Economics* 29(3): 501-516.

Flagg L.A., B. Sen, M.L Kilgore, and J.L. Locher. 2014. The Influence of Gender, Age, Education, and Household Size on Meal Preparation and Food Shopping Responsibilities. *Public Health Nutr* 17(9): 2061-2070.

Fox, J.A, J.F. Shogren, D.J. Hayes, and J.B. Kliebenstein. 1998. CVM-X: Calibrating Contingent Values with Experimental Auction Markets. *American Journal of Agricultural Economics* 80(3): 455-465.

Gracia, A., and T. de-Magistris. 2013. Preferences for Lamb Meat: A Choice Experiment for Spanish Consumers. *Meat Science* 95(2): 396-402.

Grebitus, C., H.H. Jensen, and J. Roosen. 2013. US and German Consumer Preferences for Ground Beef Packaged Under a Modified Atmosphere - Different Regulations, Different Behavior *Food Policy* 40:109-118.

Grebitus, C., H.H. Jensen, J. Roosen, and J.G. Sebranek. 2013b.Fresh Meat Packaging: Consumer Acceptance of Modified Atmosphere Packaging Including Carbon Monoxide. *Journal of Food Protection* 76(1): 99-107.

Greene, B.E., I.M. Hsin, and M.W. Zipser. 1971. Retardation of Oxidation Color Changes in Raw Ground Beef. *Journal of Food Science* 36: 941-942.

Heath, T.B., and S. Chatterjee. 1995. Asymmetric Decoy Effects on Lower-Quality versus Higher-Quality Brands: meta-Analytic and Experimental Evidence. *Journal of Consumer Research* 22(3): 268-284.

Heerwegh, D., and G Loosveldt. 2008. Face-to-Face Versus Web Surveying in a High-Internet-Coverage Population: Differences in Response Quality. *Public Opinion Quarte*rly 72(5): 836-846.

Hermel, S.R. 1993. Extending the Bloom. *Beef* 1993:8-12.

Killinger, K.M., C.R. Calkins, W.J. Umberger, D.M. Feuz, and K.M. Eskridge. 2004. Consumer Visual Preference and Value for Beef Steaks Differing in Marbling Level and Color. *Journal of Animal Science* 82(11):3288-3293.

Krosnick, J.A. 1991. Response Strategies for Coping with the Cognitive Demands of Attitude. *Applied Cognitive Psychology* (1991):213-236.

Lancaster, K. 1966. A New Approach to Consumer Theory. *Journal of Political Economy* 74: 132-157.

Levav, J., R. Kivetz, and C.K. Cho. 2010. Motivational Compatibility and Choice Conflict. *Journal of Consumer Research* 37(3): 429-442.

Lim, K.H., W. Hu, L.J. Maynard, and E. Goddard. 2013. U.S. Consumers' Preference and Willingness to Pay for Country-of-Origin-Labeled Beef Steak and Food Safety Enhancements. *Canadian Journal of Agricultural Economics* 61(1): 93-118.

Litzenberg, K.K. and V.E. Schneider. 1987. Competencies and Qualities of Agricultural Economics Graduates Sought by Agribusiness Employers. *American Journal of Agricultural Economics* 69(5): 1031-1036.

Loomis, John B. 2014. Strategies for Overcoming Hypothetical Bias in State Preference Surveys. *Journal of Agricultural and Resource Economics* 39(1):34-46.

Loureiro, M.L., and W.J. Umberger. 2003. Estimating Consumer Willingness-to-Pay for Country-of-Origin Labeling. *Journal of Agricultural and Resource Economics* 28(2): 287-301.

Lusk, J.L. and D. Hudson. 2004. Willingness-to-Pay Estimates and Their Relevance to Agribusiness Decision Making. *Review of Agricultural Economics* 26(2): 152-169.

Lusk, J.L. 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.

Lusk, J.L., J.A. Fox, T.C. Schroeder, J. Mintert, and M. Koohmaraie. 2001. In-Store Valuation of Steak Tenderness. *American Journal of Agricultural Economics*, 83(3): 539-550.

Malhotra, N. 2008. Completion Time and Response Order Effects in Web Surveys. *Public Opinion Quarterly* 72(5): 914-934.

McFadden, D.L. 1973. Conditional Logit Analysis of Qualitative Choice Behavior. *In Frontiers in Econometrics*, edited by P. Zarembka, New York: Academic Press, 105-142.

Morrison, M., and T.C. Brown. 2009. Testing the Effectiveness of Certainty Scales, Cheap Talk, and Dissonance-Minimization in Reducing Hypothetical Bias in Contingent Valuation Studies. *Environmental Resource Economics* 44:307-326.

Noel, J., and E. Qenani. 2013. New age, new learners, new skills: What skills do agribusiness graduates need to succeed in the knowledge economy? *International Food and Agribusiness Management Review* 16(3): 17-36.

Norwood, F.B. and S.R. Henneberry. 2006. Show Me the Money! The Value of College Graduate Attributes as Expressed by Employers and Perceived by Students. *American Journal of Agricultural Economics* 88(2): 484-498.

Oppenheimer, D. M., T. Meyvis, and N. Davidenko. 2009. Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power. *Journal of Experimental Social Psychology* 45(4):867-872.

Penn, J. M., and W. Hu. 2018. Understanding Hypothetical Bias: An Enhanced Meta-Analysis. *American Journal of Agricultural Economics* 100(4): 1186-1206.

Qualtrics. 2017. Provo, Utah: Available online at https://www.qualtrics.com/ [Accessed October, 2017].

Ramanathan, R., and R.A. Mancini. 2018. Role of Mitochondria in Beef Color-A Review. *Meat and Muscle Biology* 2(1), 309-320.

Ratneshwar, S., A.D. Shocker, and D.W. Steward. 1987. Toward Understanding the Attraction Effect: The Implications of Product Stimulus Meaningfulness and Familiarity. *Journal of Consumer Research* 13(4): 520-533.

Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 82(1): 34-55.

Schläpfer, F., and B. Fischhoff. 2012. Task Familiarity and Contextual Cues Predict Hypothetical Bias in a Meta-Analysis of Stated Preference Studies. *Ecological Economics* 81:44-47.

Smith, G.C., J.B. Morgan, and J.D. Tatum. 1993. Marketing Vitamin E for Meat-Quality Enhancement in the U.S.A. *Roche Vitamin E Workshop.* Luzer, Switzerland, 1993:1-7.

Smith, G.C., J.B. Morgan, J.N. Sofos, and J.D. Tatum. 1996. Supplemental Vitamin E in Beef Cattle Diets to Improve Shelf-life of Beef. *Animal Feed Science Technology* 59(1-3): 207-214.

Smith, G.C., K.E. Belk, J.N. Sofos, J.D. Tatum, and S.N. Williams. 2000. Economic Implications of Improved Color Stability in Beef. *Antioxidants in Muscle Foods, Nutritional Strategies to Improve Quality* (2000):397-426.

Suleman, F. 2016. Employability Skills of Higher Education Graduates: Little Consensus on a Much-Discussed Subject. *Procedia-Social and Behavioral Sciences* 228:169-174.

Suman, S.P., M.C. Hunt, M.N. Nair, and G. Rentfrow. 2014. Improving Beef Color Stability: Practical Strategies and Underlying Mechanisms *Meat Science* 9:490-504.

Survey Sampling International, LLC. 2017. Shelton, Connecticut: Research Now. Accessed October 2017. https://www.surveysampling.com/.

Tonsor, G. T., T.C. Schroeder, J.A. Fox, and A. Biere. 2005. European Preferences for Beef Steak Attributes. *Journal of Agricultural and Resource Economics* 30(2): 367-380.

Turnbull, B.W. 1974. Nonparametric Estimation of a Survivorship Function with Doubly Censored Data. *Journal of the American Statistical Association* 69(345):169-173.

United States Census Bureau. 2018. QuickFacts, Washington, DC: U.S. Department of Commerce. Available at https://www.census.gov/quickfacts/fact/table/US/PST045217

Velasco, M.S. 2012. More than Just Good Grades: Candidates' Perceptions About the Skills and Attributes Employers Seek in New Graduates. *Journal of Business Economics and Management* 13(3): 499-517.

Williams, S.N., T.M. Frye, M. Frigg, D.M. Schaefer, Scheller. K.K., and Q. Liu. 1992. Vitamin E as an In Situ Postmortem Pigment and Lipid Stabilizer in Beef. *Proc. Annual Pacific Northwest Animal Nutrition Conference.* Spokane, WA:149-172.

APPENDICES

Oklahoma State University Institutional Review Board

Date:	Friday, August 4, 2017
IRB Application No	AG1737
Proposal Title:	Steak and Ground Beef Discoloration Choice Experiment
Reviewed and Processed as:	Exempt

Status Recommended by Reviewer(s): Approved Protocol Expires: 8/3/2020

Principal Investigator(s):

Ryan Feuz
Stillwater, OK 74078

F. Bailey Norwood
426 Ag Hall
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

☐ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval. Protocol modifications requiring approval may include changes to the title, PI advisor, funding status or sponsor, subject population composition or size, recruitment, inclusion/exclusion criteria, research site, research procedures and consent/assent process or forms.
2Submit a request for continuation if the study extends beyond the approval period. This continuation must receive IRB review and approval before the research can continue.
3Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of the research; and
4Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Dawnett Watkins 219 Scott Hall (phone: 405-744-5700, dawnett.watkins@okstate.edu).

Sincerely,

Hugh Crethar, Chair
Institutional Review Board

Appendix 1A. IRB Approval Letter Essay 1

Appendix 2A contains demographic summary statistics for *study 1*. By comparing data from the United State Census Bureau (2018) it can be seen that overall the sample 'gender' and 'income' demographics are similar to the U.S. population. In 2017 the U.S. was estimated as 50.8% female with a median income of $55,322. Age and education attainment in the sample may be slightly misrepresented. In 2017, 87% of the U.S. population (age≥25) had an education level of high school graduate or higher while 30.3% had a Bachelor's degree or higher. This would indicate that our sample has a slightly higher average education level than the estimated U.S. population. The median age of the US population is 37.7 which is consistent with our sample. However, our sample appears to have a slightly greater percentage of those in the 18-25 age group as well as the 50-65 group.

Appendix 2A. Summary Statistics of Survey Respondents

Description	Category	Percentage
Gender	Female	51.98%
	Male	48.02%
Age	18 to 25 years old	17.72%
	26 to 34 years old	15.41%
	35 to 49 years old†	23.72%
	50 to 65 years old	25.95%
	more than 65 years old	17.21%
Education	less than high school	1.25%
	high school/GED	18.31%
	some college	24.19%
	2-year college degree†	11.25%
	4-year college degree	29.95%
	graduate degree	15.05%
Income	less than $30,000	26.93%
	$30,000 to $59,000†	27.75%
	$60,000 to $89,999	19.33%
	$90,000 to $119,999	11.84%
	$120,000 to $149,999	7.53%
	$150,000 or more	6.62%
Shopper	Is the primary shopper in the household	88.51%
	Is not the primary shopper in the household	11.49%
Frequency of steak purchase	never	9.17%
	less than once a month	22.74%
	about once a month	23.48%
	about once every two weeks	22.89%
	about once a week	16.35%
	more than once a week	5.37%
Frequency of ground beef purchase	never	6.12%
	less than once a month	14.11%
	about once a month	20.85%
	about once every two weeks	27.52%
	about once a week	24.54%
	more than once a week	6.86%

Note: † Indicates median category.

Survey Sheet 1

Carefully observe the three steaks in front of you. Pick the steak that you like the LEAST and indicate the extent to which you agree or disagree with the following statements about the chosen steak.

	Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree nor Disagree	Somewhat Agree	Agree	Strongly Agree
(1) The quality of This ribeye steak is high.	☐	☐	☐	☐	☐	☐	☐
(2) I would seriously consider purchasing this steak.	☐	☐	☐	☐	☐	☐	☐
(3) If offered at a reasonable price, people would purchase this steak.	☐	☐	☐	☐	☐	☐	☐
(4) Most people would like the taste of this steak.	☐	☐	☐	☐	☐	☐	☐
(5) I would purchase this steak for $11.25 per pound.	☐	☐	☐	☐	☐	☐	☐
(6) I would purchase this steak for $8.50 per pound	☐	☐	☐	☐	☐	☐	☐

Appendix 2B. Example Survey Statement and Response Format

Appendix 2C Results of Combined Statements 5 and 6 Two class Finite Mixture Model

	Class 1		Class 2	
	Coefficient	P-value	Coefficient	P-value
Intercept	7.5659	0.000	7.6624	0.000
Discolor	0.1610	0.558	0.0211	0.904
Price	-0.3491	0.000	-0.1640	0.000
Probability of Respondent Belonging to Class	54.21%		45.80%	

AIC:813.69 BIC:844.78

Oklahoma State University Institutional Review Board

Date:	Thursday, December 7, 2017
IRB Application No	AG1762
Proposal Title:	Steak Discoloration Table Experiment
Reviewed and Processed as:	Exempt

Status Recommended by Reviewer(s): Approved Protocol Expires: 12/6/2020

Principal Investigator(s):

Ryan Feuz
Stillwater, OK 74078

F. Bailey Norwood
426 Ag Hall
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

☐ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

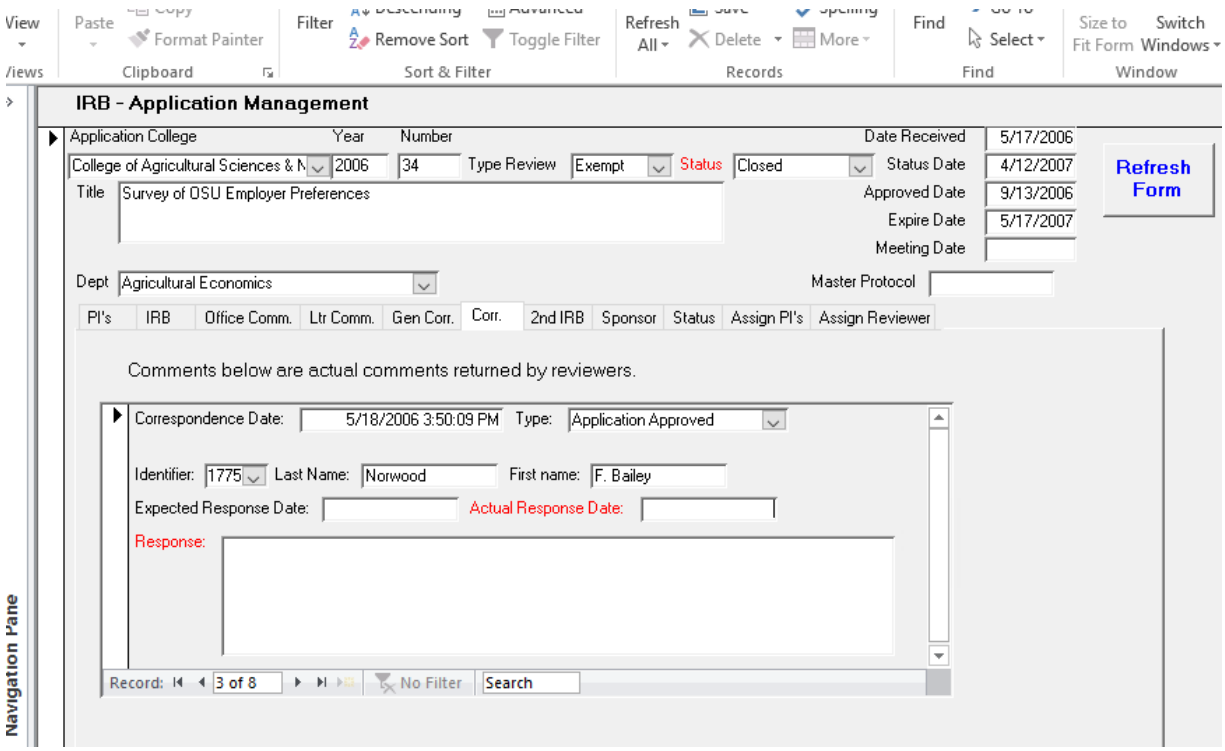
1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval. Protocol modifications requiring approval may include changes to the title, PI advisor, funding status or sponsor, subject population composition or size, recruitment, inclusion/exclusion criteria, research site, research procedures and consent/assent process or forms.
2. Submit a request for continuation if the study extends beyond the approval period. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of the research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Dawnett Watkins 219 Scott Hall (phone: 405-744-5700, dawnett.watkins@okstate.edu).

Sincerely,

Hugh Crethar, Chair
Institutional Review Board

Appendix 2D. IRB Approval Letter Essay 2



Appendix 3A. IRB Approval Essay 3- Screen Shot from IRB Office Documenting Approval

The survey for essay three was conducted in 2006. All of the necessary approval was sought and granted from the Oklahoma State University Institutional Review Board. Due to the age of the survey no letter of approval was able to be furnished by the IRB office. In place of an approval letter the IRB office provided the screenshot pictured in Appendix 3A above that shows the study was approved. The associated approval number is AG 06 34.

VITA

Ryan Feuz

Candidate for the Degree of

Doctor of Philosophy

Thesis: THREE ESSAYS: VALUATION FOR AND MARKETING SPILLOVER OF DISCOLORED BEEF AND EMPLOYER WILLINGNESS TO PAY FOR RECENT COLLEGE GRADUATE ATTRIBUTES

Major Field: Agricultural Economics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in July, 2019.

Completed the requirements for the Master of Science in Applied Economics at Utah State University, Logan, Utah in 2016.

Completed the requirements for the International Food and Agribusiness Master of Business Administration at The Royal Agricultural University, Cirencester, UK in 2017.

Completed the requirements for the Bachelor of Science in Economics and Finance at Utah State University, Logan, Utah in 2014.

Experience:

Oklahoma State University, Department of Agricultural Economics
Graduate Research Jan. 2017 – July 2019
Instructor, Quantitative Methods for Agricultural Economics Spring 2019

Northwest Farm Credit Services, Moses Lake, WA May 2014-Aug. 2015
Credit Officer