Marketing for Sustainability:

Government Management of Wild Horses and Producer Date Labeling of Foods

Dissertation

Presented in Partial Fulfillment of the Requirements for the Doctor of Philosophy in the Graduate School of The Ohio State University

By

Kathryn Elizabeth Bender

Graduate Program in Agricultural, Environmental, and Development Economics

The Ohio State University

2019

Dissertation Committee

Brian Roe, Advisor

Wuyang Hu

Zoë Plakias

Copyrighted by

Kathryn Elizabeth Bender

2019

Abstract

An ecologically unsustainable number of wild horses and burros are roaming the rangelands in the western United States. Options for dealing with this excess are limited as wild horses and burros are federally protected by a 1971 Congressional Act. The Bureau of Land Management is tasked with managing these horses. One strategy is to gather horses from the range and care for them in off-range facilities. In some cases, these captured horses are placed into private homes through online auctions. Data from these auctions allow analysis of the market for wild horses and burros. In my first essay, I utilize this data to determine the market value of characteristics of wild horses to deduce the effect of physical characteristics and training on the sale of wild horses. I find that specific traits such as color, height, and birthplace impact the sale value while saddle training greatly increases both the probability of sale and sale value. These results can help inform decisions made on the management of these horses going forward not only through sales but through management of the quality of stock on and off the range.

In my second essay, I attempt to determine the value of characteristics of a different kind of good: information on food quality. Sensory evaluation has been used to detect consumer behavior in response to label characteristics. I combine sensory and choice experiments to analyze how label type and freshness indicated on packages affect consumer purchasing behavior. The study uses data from a discrete choice experiment to estimate consumer preference for milk package attributes such as product freshness and

type of label used to communicate freshness. The results from a mixed logit estimation show that the average consumer is willing to pay $0.20 more per carton for each additional day of shelf life indicated on the label, regardless of the type of label employed. However, consumers with low discard sensitivity are willing to pay less for additional shelf life than those with average and high discard sensitivity. Furthermore, the results show that the preference for the standard date label is stronger for older consumers than younger consumers with older consumers willing to pay $0.56 more for the date label. The study recommends that producers implementing food waste reduction strategies consider the impact on each type of consumer to minimize food waste while maximizing profits.

In my third essay, I build on the results of the second essay. In this study, I examine not only the effect of the date but also the effect of the phrase that accompanies the date on consumer discard behavior. This experiment expands to four products in addition to milk: fresh chicken, lettuce, bread and cereal. I employ a logit model to determine the effect of phrase-date combinations and information about date labeling on consumer discard behavior and safe consumer management of food. I find that the difference between two date labeling schemes, the status quo labeling and the two-phrase standardized labeling, is often insignificant while the main effect of labeling on discard is driven not by the phrase but rather the date. This is an important finding for manufacturers and public policy makers; if manufacturers switch to standardized phrases and extend the date on the label to agree with the standardized definition, unknowing consumers may continue to ignore the phrase unless an intense information campaign accompanies the change. Manufacturers could be conservative extending the date on the label, but this would counter the point of a

standardized label system to reduce consumer confusion. Nevertheless, any policy that is implemented by the government, nonprofits, or manufacturers needs to consider the unintended consequences that may occur as evidenced in this essay.

Dedication

To Grandma and Grandpa

Acknowledgments

I would like to express my appreciation for my advisor, Brian Roe, who has guided me through the doctoral program over the last five years. His support and guidance were essential to my success throughout graduate school. I would also like to thank my candidacy and dissertation committee members, Wuyang Hu, Zoë Plakias, and Sathya Gopalakrishnan, as well as Jill Stowe, for their valuable feedback.

My classmates were vital to my success as well, and I am grateful for the support and friendship of Shang Xu, Corinne Bocci, John Dougherty, Jian Chen, Beth Robison, Chris Dunphy, Khushbu Mishra, Wade Litt, Rodrigo Perez, and Tony Gallenstein. Aishwarya Badiger worked with me on several projects and is a great research partner and friend. I am also indebted to Diane Ross for her friendship and mentorship throughout my time in graduate school.

Thank you to my colleague, friend, and sister, Rosemary Emegu Isoto, to whom I will forever be indebted for her valuable advice, feedback, and tough love. I am so grateful to have you in my life.

Finally, I would like to thank my family for their never-ending support of my many adventures. I love you Mom, Dad, Grandma, Bob, Jake, Abby, Lacey, and Luke. Mom and Dad, none of this would have been possible without you. From the bottom of my heart, thank you.

Vita

Education

2008 ..... Big Walnut High School

Sunbury, Ohio

2011 ..... Centre College

B.S. Economics

Danville, Kentucky

2016 ..... The Ohio State University

M.Sc. AED Economics

Columbus, Ohio

Work Experience

2015 – 2019 ..... The Ohio State University

Research and Teaching Assistant

Columbus, Ohio

2017 – 2018 ..... Otterbein University

Adjunct Professor

Westerville, Ohio

Publications

Roe, Brian E., Phinney, David M., Simons, Christopher T., Badiger, Aishwarya S., Bender, Kathryn E., and Heldman, Dennis R. (2018). *Discard Intentions are Lower for*

Milk Presented in Containers without Date Labels. Food Quality and Preference 66, 13-18. DOI: 10.1016/j.foodqual.2017.12.016

Roe, Brian E. & Bender, Kathryn. (October 2016). *A Cost-Benefit Analysis of the 4R Nutrient Stewardship Certification Program*. Available online at https://4rcertified.org/wp-content/uploads/2017/02/4R-Cost-Benefit-Analysis-10.31.16.pdf

Fields of Study

Major Field: Agricultural, Environmental, and Development Economics

Table of Contents

List of Tables

List of Figures

Chapter 1.

Adoption of Wild Horses: Evidence from Online Auctions of Wild Horses in the United States

1.1. Introduction

Wild horses and burros have roamed the western United States since the 16th century. Cultures have revolved around the partnership between man and horse while others have loathed the wild horses and the loss of resources brought about by herds. As a nation, the United States is struggling to balance the interests of people from both sides, protecting the horses as part of the national heritage while considering the burden placed on tax payers and local communities and ecosystems. This struggle has resulted in an excess number of wild horses and burros on the range and in government holding facilities, which costs the government millions of dollars every year with no apparent relief over the next couple of decades (U.S Department of Interior, 2016).

The wild horses that reside in the western United States originate from Spanish conquistadors who brought the horse during expeditions to explore the "New World." Over the years, herds of feral horses formed as domesticated horses were set loose, escaped, or stolen in raids. By the mid-1800s, there were an estimated two million wild horses roaming the plains. These feral horses competed with domesticated horses and cattle for grass,

leading to many questionable methods of disposing of them.[1] Even after legislation was passed in 1959 outlawing the use of motorized vehicles to gather wild horses, these questionable disposal methods continued, and by 1970, only an estimated 18,000 were left (De Steiguer, 2011).

In 1971, after public outcry over the plight of the wild horse, Congress deemed wild horses and burros to be "living symbols of the historic and pioneer spirit of the West" and "an integral part of the natural system of the public lands."[2] This turned the wild horses and burros from being considered a feral animal with an open hunting season to a federally protected animal. The Bureau of Land Management (BLM) within the Department of the Interior was tasked with managing the numbers and health of wild horses and burros.

The BLM balances the numbers with the needs of other animals on the land and the number of grazing permits given to cattle ranchers. Outside of foaling season, the BLM conducts roundups, otherwise known as "gathers," to help keep the population under control. With no natural predators or fertility control program, herds can double in size every four years[3] creating a growing problem for the BLM (U.S. Department of Interior, 2016).

---

[1] Many were sent to slaughter houses to be turned into fertilizer or pet food. Others were crippled or killed outright by gunshot or tied up and left to die. Still others were run to exhaustion by motorized vehicles or helicopters. During the World Wars, some were rounded up and sent to Europe to help in the wars.

[2] *"Congress finds and declares that wild free-roaming horses and burros are living symbols of the historic and pioneer spirit of the West; that they contribute to the diversity of life forms within the Nation and enrich the lives of the American people; and that these horses and burros are fast disappearing from the American scene. It is the policy of Congress that wild free-roaming horses and burros shall be protected from capture, branding, harassment, or death; and to accomplish this they are to be considered in the area where presently found, as an integral part of the natural system of the public lands.*" -The Free-Roaming Wild Horses and Burros Act of 1971

[3] Chemical vasectomy and two promising immunocontraceptives have been identified as the most promising fertility control methods, but more research needs to be done before implementing these methods.

Wild horses and burros are managed on herd management areas (HMAs) that span 31.6 million acres across the western United States. Figure 1 is a map of the current HMAs managed by the BLM and is color coded by the equines managed in that area: horses, burros, or both. As of March 2019, there were 88,090 wild horses and burros on these HMAs, which far exceeds the 26,690 that can be sustained on those lands. In addition, the BLM cares for 50,020 horses and burros in off-range holding facilities, a number that is fast approaching the total capacity of 56,106. Some of these horses are in short-term facilities with a chance of being adopted while others are held in long-term facilities with no chance of adoption (BLM, 2019). In 2008, the average cost of caring for a horse in short-term and long-term facilities was $5.08 and $1.27 per day, respectively (GAO, 2008). While the cost of short-term care has not changed significantly, it is estimated that the cost of long-term care has increased to around $2.00 per horse per day. This means that, on average, a young horse taken off the range and kept in a holding facility for the remainder of its life will cost the government more than $19,000.[4]

[4] On average, a horse will stay for 210 days in a short-term holding facility and up to 25 years in a long-term holding facility.

Figure 1. Herd Management Areas in the Western United States

Wild horses and burros are initially rounded up during "gathers" that occur at the BLM's discretion. The BLM does not conduct gathers from March through June, which is foaling season during which pregnant females give birth. During a gather, herds are driven off the range and into holding pens where those that are selected for removal are separated from those that are released. Removal is based on the age and adoptability of the horse. Those that are under five years of age are the first to be removed followed by those over

ten. Males (stallions and colts) that are selected for removal are castrated (gelded)[5] so that there is no risk of the removed horses reproducing once in captivity. Females (mares and fillies) selected for removal are often pregnant at the time of their capture and give birth in captivity.

The BLM uses private adoption to manage the excess of wild horses. Since 1971, the BLM has facilitated adoption of 240,000 wild horses (and burros) through off-range corrals, adoption events throughout the country, and online auctions (BLM, 2019). Currently, the minimum adoption fee is $25, adopters must be approved by the agency, and approved adopters can purchase no more than four equines every six months. In addition, the title to the equine is not transferred to the new owner for one year.[6] These restrictions come after public outcry about a much less restrictive 2018 policy that allowed any individual to adopt 24 equines per day with an immediate title transfer. Wild horse advocates viewed this as an attempt by the BLM to decrease the number of horses in its care through making the horses available to "kill buyers." While it is illegal for an equine to be adopted for the purpose of slaughter, the enforcement of this requirement is difficult. The continued public outcry over any actual or perceived mismanagement of wild horses and burros is a testament to value the public places on them. This non-use value is only one of the benefits individuals derive from wild horses.

Adoptions demonstrate the use value that individuals place on wild horses and burros. At most adoption events throughout the country, private individuals can adopt an

[5] Gelding is the process of castrating a male horse. Once gelded, these horses are referred to as "geldings."
[6] This is meant to prevent kill buyers from adopting wild horses to take to slaughter houses abroad.

equine for a flat adoption fee on a first-come, first-served basis. Online auctions are also held approximately once every other month and are a more cost-efficient way to adopt out a horse than the satellite adoption events since the horse can remain at the short-term holding facility rather than be shipped to event locations; however, there is still a cost to entering horses in the online auction. Pictures along with a description of the physical characteristics and training stage of each horses are posted for viewing in advance of the auction. For an often-short-staffed BLM, this time commitment is significant, so it is not feasible to enter all horses residing at short-term facilities in each auction.[7] The auctions run for one to two weeks with individuals being able to place bids throughout that time. At the end of the process, the winner can pick up the horse from its current location or choose to have the horse brought to another location or adoption event closer to the adopter's location. The title for the horse is transferred to the buyer after one year of ownership (BLM, 2019).

While the rate of adoption is low compared to the rate of wild horses captured by the BLM – for example, in the fiscal year 2018 there were 3,158 adoptions to 11,472 captured off the range – adoption is the main way the BLM can avoid the cost of caring for the animal. In this essay, I explore the role of private adoption in the BLM's herd management policy using data from online auctions to determine the value of different characteristics of wild horses as well as the benefits of training. I apply a hedonic pricing model to the American wild horse market in the United State as well as analyze

[7] Records on the result of the auctions are rarely saved due to the limited BLM staff available.

characteristics that increase the probability a horse will sell, thus allowing the government to forgo the cost of caring for the animal for the rest of its life.

1.2. Background of the Equine Market in the United States

Much of the current literature has focused on the management of wild horses and proper methods to control growth (Mask, Schoenecker, Kane, Ransom, & Bruemmer, 2015; Swegen and Aitken, 2014; Bechert, Bartell, Kutzzler, Menino, Bildfell, Anderson, & Fraker, 2013; Kirkpatrick, Rutberg, Coates-Markle, & Fazio, 2012; Ransom, Roelle, Cade, Coates-Markle, & Kane, 2011). Only a few papers cover the characteristics of wild horses that increase their chances of being adopted; however, none of these papers examines the market after both the 2008 recession and the elimination of horse slaughter in the United States as this essay does.

Lenz (2009) discusses the problem of unwanted horses in the United States. In 2007, slaughter was made illegal in the two states in which there were slaughter houses for horses. Data from the USDA showed that the majority of horses sent to slaughter were riding horses, and less than one percent were BLM wild horses. However, there is still a degree of substitutability between riding horses and mustangs so the ban on slaughter affects the market for mustangs indirectly through prices for riding horses (Koncel and Rutberg, 2012). In 2005, at an American Association of Equine Practitioners meeting, it was agreed that the horse industry was not prepared to handle the excess of unwanted horses that were going through processing plants at the time. This excess supply of grade

horses[8] would be in competition with BLM's wild horses (Ahern, 2006). Furthermore, without a bottom on the market from slaughter prices, the prices of grade horses could be driven toward zero (North, Bailey, & Ward, 2005).

The effect from the 2007 slaughter ban in the United States was followed by the recession in 2008, which had a significant impact on the horse industry. The recession confounded the effect of the slaughter ban on prices in the horse market making it difficult to distinguish the effect from slaughter and the effect from the recession. However, a study by the Government Accountability Office (2011) determined that the slaughter ban did have a significant negative effect on prices of horses — although this is still disputed by many groups. This essay uses data generated after both of these events, yielding a more accurate portrayal of the current equine market than studies done before or during 2008.

Several papers have analyzed the sale prices of horses, but most examined racehorses which have different uses than wild horses (Maynard and Stoeppel, 2007; Neibergs, 2001; Vickner & Koch, 2001; Lansford, Freeman, Topliff, & Walker, 1998). Stowe and Kibler (2016) model the characteristics of Thoroughbred racehorses at nonprofit organizations dedicated to placing retired racehorses[9] into new homes. They find that the time it takes a horse to be adopted is determined by gender, color, age, and most significantly, soundness.[10]

[8] A grade horse is one of unknown parentage or mixed breeding. Most wild horses would be categorized as such.

[9] Thoroughbred racehorses are retired after relatively short racing careers. Many are successful in second careers, but just like wild horses, a significant amount of time goes into training them after they come off the track.

[10] Soundness represents the health of the horse. A horse with a limp or breathing issues would be considered unsound, while a healthy horse that safely could be used for riding would be considered sound.

An important aspect of the BLM wild horse adoption process is the stipulations put on the new owners. Specific transport means and fencing are required before owners can take their horses home. Horses who have not been trained to tie need to be transported in a stock trailer for the safety of the horse. Many potential adopters modify their barns and pastures to make them safe for a wild horse. Due to most of the horses available for adoption being wilder at first than a standard domesticated horse, these modifications are necessary to make sure a wild horse can be safely contained within the barn space. In addition, personnel that deal with wild horses need to be trained properly, which could lead to higher costs of ownership. For example, some farriers[11] won't work on wild horses and in those cases, the owner must find a specialized farrier who would be willing to take them on as a client. These farriers are likely to charge more based on the added risks of working with a wild horse. In addition to these things, the title is not transferred to the owner for one year to ensure the horse's protection for that year. These conditions and modifications necessary for owning a wild horse increase the cost of adoption (Godfrey and Lawson, 1986). A survey conducted by Koncel and Rutberg (2012) indicated that owners are not always aware of the long-term costs associated with wild horses at the time of adoption. Most of the participants in this survey had adopted from satellite adoption events.[12] This study uses adoptions from internet auctions which, unlike many of the satellite adoption events, take more than a couple of days to complete the transaction. Potential adopters must register to bid, complete an application for adoption, and then participate in the two-week-

[11] A farrier is the person who trims horses' hooves, a necessary part of care for horses in captivity that should be done every 6-8 weeks.

[12] The BLM holds satellite adoption events across the country where people can adopt a horse at the event and take it home that weekend.

long auction before making arrangements to pick up their horse. This most likely decreases the probability that a sale from an internet auction is an "impulse buy" with the adopter being unaware of the true costs associated with a wild horse as is sometimes the case at adoption events (Koncel & Rutberg, 2012).

Elizondo, Fitzgerald, and Rucker (2016) study all wild horses adopted through the BLM's Adopt-a-Horse Program[13] from 1997 to 2010. The authors found that costly horse training reduced BLM's total costs. Specifically, horses with training[14] brought in $120 more than the sum of the training costs paid by the BLM and the average fee paid for adopted horses without training. In their sample, all horses that received training were adopted so training perfectly predicted the probability of being adopted. While training might have led to the guaranteed sale of horses before 2010, this is no longer true based on more recent data. Whether this change is due to a decrease in demand, scaling up of the training program, or some other factor is undetermined. However, the training program has been scaled up significantly with 1,754 equines trained in 2017 versus 667 in 2012 (BLM, 2019). Elizondo et al. (2016) also examined physical characteristics of the horses offered for adoption and eventually adopted. They found that mares and horses with unique coat colors were more likely to be adopted out than geldings and horses with basic coat color. Limitations of their study is that the horses' training level is not broken down by the type of training and the time period studied. The study ended in 2010 when the effects of the recession were still being felt through the equine market. The demand for wild horses

[13] Adopt-a-Horse Program includes internet auctions and satellite adoption events.
[14] The exact level of training is not specified (e.g., halter-broke, saddle-broke, etc.).

continued to change after the period of the Elizondo et al. (2016) study as the market adjusted to the rippling effects of the recession.

This essay examines the effect of physical characteristics and training on the private sale of wild horses in the post-recession equine market. As the number of excess horses in the BLM's care increases, an understanding of the potential for private sales to alleviate the BLM's costs is vital to determine the optimal policy solution. This will be a solution to a dynamic model where the physical characteristics of the horses left on the range, given fertility control, brought up in a gather, or made available at adoption events are part of the optimal policy solution. The stock of the horses on the range and reproducing will determine the supply of wild horses from which some will be captured during a gather and made available for adoption. Therefore, the future quality of the stock on the range needs to be considered alongside the present need to move horses into private individuals' care. In addition, physical characteristics will need to be considered when implementing any fertility control methods with horses left on the range. This essay takes the first step in finding an optimal solution by analyzing the impact of these physical characteristics on private adoption. Training is also analyzed as a potential method to increase the number of horses going to private homes.

1.3. Data, Variables, and Summary Statistics

1.3.1. Data

The data were hand collected from the BLM Internet Adoption Program's online gallery from November 2012 through November 2014. Each wild horse selected for the internet auction is given a webpage posted before the auction with basic information

including color, height, gender, age, and at least one picture. If the horse is trained, a description of the training may appear on the page as well. The pages are grouped together based on location of the horse. In many cases, the horse can be shipped to different locations for pickup at the buyer's request. Auctions are English style with bidding beginning at $125 and increasing by a minimum of $5 with each bid. The auction is open to bidding for two weeks at which time the bidder with the highest bid wins and pays the amount of their winning bid.

Initially, there were 1,524 observations from the horses gathered and put in the online auction. Burros were dropped since there are few burros available and the market for burros is significantly different than the market for horses. There were also a few stud colts that were too young to geld and were removed from the sample. Horses with missing information on height, gender, and capture date were also dropped leaving a total of 1,256 observations. While the majority of horses only appear once in the dataset, some horses appear multiple times if they failed to sell the first time they were put up at auction. On average, horses went to auction twice although one horse appeared at auction nine times and several others were entered eight times.

Table 1 presents the summary statistics of all horses entered in online auctions and the subsample of horses that sold in online auctions. The average age was 3 years with the youngest horse being six months and the oldest being 14 years. The ages are skewed toward younger horses with 86% of the sample being between one and four years old. This is likely due to the policy of prioritizing the removal of horses under five from herds during gathers.

Horses range in height from 11 to 16.1 hands[15] with an average height of about 14.1 hands (57 inches).

With 630 sales and 626 failures, just over half of the internet auction entries were sold. The minimum bid for online auctions is $125. This was the lowest price for which any horse sold with 45.2% of horses sold selling at this price. Among the horses that sold, 90% sold for less than $500. An untrained, 3-year-old pinto mare sold for the highest bid of $2,595. The average winning bid among the horses that sold was $233.92.

The level of training is a three-level variable: no training, halter broke, or saddle broke. Training progresses from a horse learning to give to a halter to accepting a saddle and rider on its back; therefore, all horses that are broke to saddle must be broke to halter first. A total of 93 horses (7.5%) in the sample are halter broke and 41 (3.3%) are also saddle broke. A horse must be halter broke before being saddle broke, so there are 52 horses that are solely halter broke in the sample. Both tasks take a considerable amount of time to achieve with a wild horse, so these low numbers are to be expected.

The majority of auctions, about 93%, are held between March and November. An average of 98 horses were entered in each online auction with auctions having as few as three horses to as many as 171 horses entered. All auctions had horses that sold and horses that did not sell. The least successful auction had 38.8% of entries sell while the most successful auction had 97.2% of entries sell which was all but one entry.

[15] A "hand" is the common measure of height for horses where one hand is equal to four inches.

Table 1. Summary Statistics of Horses Entered in Online Auctions

Variable	Description	Entered	Sold
Physical characteristic			
Mare	Female horse	0.613	0.578
		(0.487)	(0.494)
Age	Age of horse at time of auction	3.060	2.980
		(1.470)	(1.530)
	Minimum age	6 months	6 months
	Maximum age	14 years	9 years
Base	Solid base coat color	0.506	0.359
		(0.500)	(0.480)
Pattern	Gray, roan or white coat color	0.210	0.221
		(0.408)	(0.415)
Dilute	Diluted coat color (palomino, buckskin, dun)	0.194	0.289
		(0.396)	(0.454)
Spotted	Spots or blanket pattern on coat color	0.019	0.021
		(0.137)	(0.142)
Pinto	Pinto coat color	0.081	0.129
		(0.273)	(0.335)
Height	Height at the withers in inches	56.80	56.95
		(3.31)	(3.64)
Bornincap	Born in captivity	0.236	0.224
		(0.425)	(0.417)
Training			
Halter	Horse is trained to lead	0.075	0.127
		(0.263)	(0.333)
Saddle	Horse is started under saddle	0.033	0.634
		(0.178)	(0.244)
Auction			
Available	Number of horses available in online auction	98	91
		(42)	(41)
Spring	March, April, or May auction	0.209	0.219
		(0.407)	(0.414)
Summer	June, July, or August auction	0.321	0.298
		(0.467)	(0.458)
Autumn	Septmeber, October, or November auction	0.393	0.383
		(0.489)	(0.486)
Winter	December, January, or February auction	0.076	0.100
		(0.266)	(0.300)
Chances	Number of times horse went to auction	2.07	1.37
		(1.70)	(0.83)
Sold	Horse successfully sold at auction	0.502	1.000
		(0.500)	0.000
Price	Winning bid of horse sold at auction	--	233.92
			(265.11)
Observations		1256	630

Mean reported unless otherwise specified

1.3.2. Empirical Methods

I first examine the factors that affect the probability a horse is sold using the following probit model.

$$\pi_{ij} = \beta_0 + \beta_1 \boldsymbol{X}_i + \beta_2 \boldsymbol{W}_j + \varepsilon_{ij} \quad (1)$$

π_{ij} is the probability of the horse *i* being sold in auction *j*; $\boldsymbol{X}_i$ is a vector of horse-specific characteristics including age, gender, color, training, captivity history, and captivity interacted with age; and $\boldsymbol{W}_j$ is a vector of auction-specific characteristics including time of year and the number of horses available in the auction.

I interact age with captivity in this model since being held in captivity could be viewed differently based on the age of the horse. There may be a balance between time spent on the range and time spent in captivity that is appealing to people. A horse that was captured young might seem to be less of a true "mustang" than one that grew up in the wild, so the effect of being born in captivity would be negative. Meanwhile, a longer period of captivity for an older horse may be viewed as beneficial to the gentling process.

Studies in the past have found gender to be significant, but analysis of more recent sales has found a lesser effect. I include it in this model as the issue is still unresolved (Elizondo et al., 2016; Godfrey & Lawson, 1986; Stowe & Kibler, 2016). I expect the later seasons of the year to have negative effects since the market for horses usually declines heading into winter. Wild horses are on the smaller side for an equine, so buyers usually like to find ones that are taller to ensure the horse's ability to carry a rider. Therefore, height is expected to have a positive effect on the probability of sale. Different colors are considered unique, so in relation to a base coat color, I expect to see positive

signs on all the color covariates. Training should also improve the probability of sale while the number of horses in the auction (*availability*) should lower the chance of being sold. Due to the saturated market, the more horses that are offered, the less likely each individual horse is to sell. *Chances* specifies the number of times are horse has been a part of an online auction. The effect of *chances* is negative as a horse that has been presented at auction multiple times is less likely to sell by definition of the variable; the horse has failed to sell for at least all but one of the times it was entered in an auction. The more horses that are offered, the less likely each individual horse is to sell. The effect of being born in captivity could be positive or negative.

The probit model is run four times with different combinations of location and year fixed effects. Location is coded as the facility where the horse is currently held. I expect both will have an effect on the probability of being sold and will need to be included. The results from this model are important to the BLM not only through increasing revenue but also foregone costs. By selling a horse, the BLM avoids the expense of caring for that horse for the rest of its natural life. Determining the characteristics that get a horse moved from the government's care into a private owner's care could allow the BLM to more efficiently decide about which horses to put up for sale and when to do so.

Next, I estimate the effects of the horse and auction characteristics on the sale price. The prices that are observed are the prices of horses that sold meaning that their value is at least the minimum fee of $125. This limitation on the sale prices yields truncated data. To account for this selection bias, I employ a Heckman model where

$$\ln(price) = \xi(\Pr(sold), \boldsymbol{Z}_i, \boldsymbol{R}_j) \quad (2)$$

and the probability of being sold is estimated with the variables discussed above; $\boldsymbol{Z}_i$ is a vector of horse-specific characteristics including age, gender, color, training, and whether the horse was born in captivity; and $\boldsymbol{R}_j$ is a vector of auction-specific characteristics including time of year and the year in which the auction took place but not the number of horses available in the auction. There are four variables included when estimating the probability of the horse selling but not included when estimating the sale price – *captivity*, *captivity*age*, *available*, and *chances* – thus satisfying the exclusion restriction.

1.4. Results

1.4.1. Probability of Sale

Table 2 shows the coefficients for different specifications of the probit model. Model 1 does not control for location of the horse or the year. Model 2 controls for the year, and Model 3 controls for the location. Model 4 controls for both year and location and is the preferred model for the remainder of the analysis due to the results discussed below.

Mares are 7.0 percentage points less likely to sell than geldings when facility and time fixed effects are included. This is a different result than what Elizondo et al. (2016) found for horses sold before 2010; in their sample, mares were more likely to sell than geldings. In this sample, the older the horse, the less likely it is to sell, but this is only true within the age range, 6 months to 14 years, of this sample. Height is significantly positive with a one-inch increase in height increasing the probability of sale by 2.9 percentage points. All the color covariates are positive and significant compared to the base coat

colors. The unique colors of a pinto and a base color with a diluted gene are the most likely colors to sell, each increasing the probability of sale by about 33 percentage points. The spotted pattern partially or fully covering the body and color pattern (gray, roan, or white) increases the probability of sale by 26.7 and 18.6 percentage points, respectively. These results indicate that a solid base coat color is the least desirable color at auction.

Captivity and captivity interacted with age are both significant with a negative and positive effect, respectively. This indicates that the length of captivity decreases the probability of sale for a young horse more than it does for an older horse. This supports the hypothesis that a horse that was captured young may seem to be less of a true mustang than one that grew up in the wild while a longer period of captivity for an older horse may be seen as beneficial to the gentling process. The number of times a horse is entered in an auction, represented by *chances*, is negative and significant as expected since these horses appear in the data set multiple times due to their failure to sell.

As anticipated horses entered in the spring auctions are the most likely to sell followed by winter auctions, summer auctions, and finally autumn auctions. The market for horses usually declines heading into winter as weather increases the costs of caring for a horse through limited pasture and decreases a main benefit, riding, of owning a horse. The number of wild horses entered in each online auction is negative and significant, with each additional horse added decreasing the probability of sale for each other horse by 0.2 percentage points.

The effects of training are positive in all four models, but the effect of halter training is insignificant once location fixed effects are added. Certain facilities have a better

reputation than others. The ones with better reputations are much more likely to have trained horses in the facility, so training could have absorbed some of the facility fixed effect when the fixed effect is not included in the model. Halter training is part of the gentling process, but a halter trained horse does not necessarily indicate a gentled horse. Saddle training is a much more involved process as it requires an immense level of trust for a horse to allow a human to sit astride it, a feeling similar to a predator attacking a horse in the wild. Therefore, the difference between an untrained horse and a halter trained horse can be minimal while the difference between an untrained horse and a saddle trained horse is significant. It is expected, then, that the effect of saddle training is greater than that of halter training. Saddle training a horse increased the probability of the horse selling by 45.1 percentage points.

Table 2. Effect of Horse and Auction Characteristics on Prob(Sale)

Dependent variable: Sold

	(1)		(2)		(3)		(4)	
	Probit	Marginal Effect	Probit	Marginal Effect	Probit	Marginal Effect	Probit	Marginal Effect
Physical Characteristic								
Mare	-0.087	-0.035	-0.128	-0.051	-0.118	-0.047	-0.175*	-0.070
	(0.088)		(0.090)		(0.097)		(0.098)	
Age	-0.182***	-0.073	-0.166***	-0.066	-0.167***	-0.067	-0.162***	-0.065
	(0.058)		(0.058)		(0.061)		(0.062)	
Color Pattern	0.374***	0.148	0.377***	0.149	0.456***	0.179	0.475***	0.186
	(0.105)		(0.106)		(0.111)		(0.112)	
Dilute	0.854***	0.322	0.851***	0.319	0.874***	0.328	0.877***	0.328
	(0.112)		(0.113)		(0.116)		(0.117)	
Spotted	0.764**	0.281	0.754**	0.277	0.708**	0.263	0.726**	0.267
	(0.320)		(0.317)		(0.329)		(0.328)	
Pinto	0.923***	0.335	0.947***	0.339	0.895***	0.326	0.935***	0.335
	(0.162)		(0.163)		(0.169)		(0.170)	
Height	0.066***	0.026	0.064***	0.025	0.075***	0.030	0.073***	0.029
	(0.014)		(0.015)		(0.019)		(0.019)	
Training								
Halter	0.447**	0.174	0.442**	0.172	0.284	0.112	0.235	0.093
	(0.206)		(0.209)		(0.294)		(0.300)	
Saddle	1.334***	0.424	1.351***	0.423	1.443***	0.442	1.532***	0.451
	(0.469)		(0.476)		(0.535)		(0.545)	
History								
Born in Captivity	-0.207**	-0.082	-0.179*	-0.071	-0.158	-0.063	-0.155	-0.062
	(0.106)		(0.106)		(0.116)		(0.117)	
Time in Captivity	-0.034***	-0.014	-0.031***	-0.012	-0.029***	-0.012	-0.025**	-0.010
	(0.009)		(0.009)		(0.010)		(0.010)	
Captivity*Age	0.006***	0.002	0.005**	0.002	0.005**	0.002	0.005**	0.002
	(0.002)		(0.002)		(0.002)		(0.002)	
Chances	-0.380***	-0.152	-0.383***	-0.153	-0.465***	-0.185	-0.482***	-0.192
	(0.038)		(0.040)		(0.043)		(0.046)	
Auction								
Spring	0.380***	0.150	0.390***	0.154	0.375***	0.148	0.388***	0.153
	(0.119)		(0.121)		(0.127)		(0.129)	
Summer	0.102	0.041	0.149	0.059	0.142	0.057	0.211*	0.084
	(0.110)		(0.111)		(0.116)		(0.117)	
Winter	0.481***	0.187	0.475***	0.184	0.375**	0.147	0.341*	0.134
	(0.173)		(0.179)		(0.186)		(0.193)	
Available	-0.004***	-0.001	-0.004***	-0.002	-0.004***	-0.002	-0.005***	-0.002
	(0.001)		(0.001)		(0.001)		(0.001)	
Year FE	No		Yes		No		Yes	
Location FE	No		No		Yes		Yes	
Psuedo R^2	0.264		0.274		0.297		0.307	
Observations	1256		1256		1256		1256	

*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Marginal effect at the mean reported.
Standard errors in parentheses.

1.4.2. Sale Price

In this section, the effect of horse and auction characteristics on sale price is analyzed. As discussed previously, a Heckman model is used to correct selection bias. The results of the Heckman model of the covariates on the *ln(price)* are displayed in Table 3. The first column is the results from the full sample while the second column is on the subsample of horses three years of age or older. Three years old is generally accepted as the age a horse is mature enough to be trained under saddle. The third column consists of horses too young to be saddle trained, and, for this reason, *saddle* is dropped from the model. Lambda in the first and third models is negative and significant indicating that unobserved factors that increase the probability of a horse selling are associated with lower sale prices.

After accounting for the selection bias, only a handful of the covariates have a significant effect on the price for which the horse is sold. Across the full sample, age has a negative, significant effect on sale price with each additional year of age reducing the sale price by 4.8 percent. Examining this effect across the two subsamples, it appears that from six months through two years, age does not affect sales price. Another interesting result that emerges once selection bias is corrected is that the coat color of the horse does not affect price for the most part. The majority of the estimates on coat colors is positive but not significant with the exception of the pinto and diluted coat patterns. Across the full sample, pintos sell for 29.4 percent more than solid base coat colors. Among mature horses, those three years of age or older, those with a diluted coat color command a 21.4 percent higher sale price. Of the three horses that sold for the most, two were pintos and one was

dilute coat color; the results still hold when the top three horses are dropped from the sample indicating that it is not just a few high-priced horses driving this effect. Elizondo et al. (2016) found that all color variations had a significant and positive effect on price. This change in effect of coat color from the sample used in that study, data for horse sold before 2010, and this sample, horse sold between 2012 and 2014, indicates a changing market for wild horses. Another interesting result that should be noted is that horses born in captivity bring a consistently lower price – about 13.6 percent – than those born on the range once other potentially correlated factors such as training are controlled.

Height and training should be analyzed across the subsamples since these groups of horses are at very distinct stages in their lives. Horses grow until they are five years old, but most of the growth is complete by the time they reach three. In addition, once a horse has reached three, the height they will mature to is more apparent. The insignificant effect of height in the under-three subsample is expected. For horses three and older, the height affects sale price as hypothesized with a one-inch increase in height leading to a 3.2 percent increase in price. Starting a wild horse under saddle is a time-consuming task but has the greatest effect on sale price, increasing it by 54.4 percent. If this amount is computed from the minimum price of $125, this equates to a $68 increase. Across the full sample and both subsamples, halter training does not have a significant effect on sale price.

Table 3. Heckman Correction Models by Maturity

Dependent Variable: ln(price)	All Ages	Age ≥ 3	Age < 3
	Heckman	Heckman	Heckman
Physical Characteristics			
Mare	-0.047	-0.090	-0.031
	(0.045)	(0.068)	(0.064)
Age	-0.048***	-0.051**	-0.068
	(0.017)	(0.025)	(0.109)
Pattern	0.015	0.056	0.003
	(0.060)	(0.082)	(0.092)
Dilute	0.092	0.214**	0.023
	(0.063)	(0.088)	(0.092)
Spotted	0.011	0.059	-0.168
	(0.152)	(0.184)	(0.267)
Pinto	0.294***	0.398***	0.273**
	(0.077)	(0.106)	(0.114)
Height	0.010	0.032**	0.001
	(0.007)	(0.013)	(0.012)
Training			
Halter	-0.081	0.308	0.015
	(0.096)	(0.201)	(0.115)
Saddle	0.926***	0.544**	--
	(0.128)	(0.214)	
History			
Bornincap	-0.136**	-0.199**	-0.140*
	(0.057)	(0.095)	(0.081)
Auction			
Spring	0.073	0.222**	-0.108
	(0.065)	(0.096)	(0.099)
Summer	-0.010	0.028	-0.100
	(0.058)	(0.082)	(0.085)
Winter	0.041	0.187*	-0.312
	(0.085)	(0.102)	(0.195)
Constant	4.797***	3.351***	5.438***
	(0.419)	(0.782)	(0.600)
Observations	1,256	746	510
Uncensored Observations	630	350	280
Rho	-0.413	-0.216	-0.511
Lambda	-0.223***	-0.115	-0.274**

*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses.

1.5. Conclusion

The BLM faces a growing problem with the wild horses roaming the western United States. Even when a cost-effective method of fertility treatment for wild horses is developed, the problem of excess horses in long-term holding facilities will remain for many years. Currently, the only way to disperse captured horses to the public is through adoption. The results above show that there are specific characteristics of horses that increase the probability of sale and the price at which they will sell. Horses with less desirable traits have market values below the minimum $125 auction bid price. One simple change that could lower the number of horses in holding facilities is to decrease the minimum auction bid price and adoption fee. However, $125 is a minimal cost in relation to the total cost of care for a horse, and the pool of potential adopters is small, so this is unlikely to have a significant and lasting impact on private adoptions. Proper management of wild horses will require the government to adopt a comprehensive, multi-faceted policy incorporating incremental changes to all parts of the management process.

The positive effect of training on adoption is evident in this sample. This result aligns with what Elizondo et al. (2016) found for the sample of horses sold before 2010; however, the effect of training on the price a horse sells for is lower in the more recent sample used in this essay. Elizondo et al. (2016) did not separate the types of training so it cannot be determined if this decrease is driven by one type of training or the other. Overall, training makes wild horse more suitable for a private home. In this analysis, halter training does not increase the probability of the horse selling nor the sale price; therefore, money budgeted for training should be invested in starting horses under saddle. Fewer horses will

receive training, but more horses will receive advanced training that greatly increases their probability of sale. If more money is budgeted for training, the BLM could decrease the number of horses in holding facilities by providing a horse that is suitable for a larger pool of potential adopters. However, a policy should not be implemented purely based on these results as the effect may be diminished if the program were scaled up.

Next to training, the color pattern of the horse most influenced the probability of the horse selling. Horses with unique coat colors or patterns are between 20 and 35 percentage points more likely to sell than a horse with a solid coat color. Coat color has an insignificant effect on the probability of a horse being either halter or saddle trained. This is another area where the BLM could adopt a more efficient strategy for training. Training horses with solid coat colors over horses that are more likely to be sold, even without training, could increase the overall number of horses sold.

The value, and thus desirability, of specific physical characteristics of horses should be considered when conducting gathers and administering fertility control. The question of which horses to remove from the wild is an inherently dynamic decision. Horses that have desirable traits are more likely to sell, decreasing the BLM's costs in the short-run. However, in the long-run, this lowers the number of horses with desirable traits in the population since captured horses are not permitted to reproduce. The lower quality of stock could increase the BLM's future costs. The results from this essay aid in the research on the optimal decision path for management so the BLM can implement effective gather, training, auction, and fertility administration policies. As shown by the differences discussed between the results of this analysis of more recent sales versus the results found

in previous literature, the value of certain characteristics can change over time, i.e. mares are less desirable compared to geldings than they were a decade ago, so robust data on the horses and surveys of the public could help the government determine the current value of these characteristics.

A limitation of this study is that the model used is static, not dynamic. Higher sales in one period are likely to drive down sales in the next period, especially with a small pool of potential adopters who are equipped to care for a wild horse. As discussed in the previous paragraph, selecting horses with desirable traits to remove from the range for adoption could increase future costs through the lower quality of reproducing stock; however, the stock of horses in long-term holding facilities is significant and currently unmanaged in terms of private transfers. Once a horse goes to a long-term holding facility, they no longer appear at adoption events or in online auctions. This unmanaged stock could be part of an initial correction by the BLM to sell desirable horses in long-term care to private individuals. Newly gathered horses whose characteristics make them unlikely to sell could be moved to the long-term facilities where the cost of care is lower than that of short-term facilities thus minimizing the stress on these horses and decreasing the cost of care, outcomes that would be heralded by both government officials and wild horse advocates alike.

The number of wild horses on the range, in short-term holding facilities, and in long-term holding facilities continues to grow under the BLM's current policies. This growth is ecologically, financially, and politically unsustainable. Without natural predators or fertility control, the number of wild horses on the range will double every four years.

Results from this study can aid the BLM in implementing an effective fertility control program, curbing the flow of horses into the BLM's care. In addition, they can also be used to strategically gather and sell optimal horses off the range and manage the horses currently in BLM facilities. These results are one part of determining a comprehensive management program that satisfies the public demand for protection of wild horses given financial and ecological constraints.

Chapter 2.

Consumer Food Waste and Date Labeling: Combining Sensory and Choice Experiments to Assess Novel and Standard Shelf-Life Indicators

2.1. Introduction

One-third of food produced for human consumption is wasted, leading to 1 trillion USD in private financial losses and at least an additional 1.6 trillion USD in environmental and societal costs per year (Food and Agricultural Organization of the United Nations, 2013). These costs manifest in the form of water scarcity, decreased biodiversity, negative health effects, and loss of livelihoods (FAO, 2013). In the United States, the average household spends up to 2,275 USD annually on food that goes uneaten (Bloom, 2010). One of the leading factors of consumer food waste is confusion over date labels. Some estimate the amount of food wasted due to confusion about the date label to be as high as 33% (Quested & Murphy, 2014). Consumers often misinterpret the date on a package as an indicator of safety. While true for some products, the majority of date labels indicate the day after which the quality of the product might decline. An even more conservative 'sell by' date found on many packages indicates the date by which a retailer should sell the product. This date is not meant to indicate quality or safety to the consumer but is often the only date that consumers see on many packages. The waste created from discarding foods

with a quality date due to safety concerns and from discarding foods with a sell by date due to quality or safety concerns is unwarranted.

While consumer preferences may dictate the optimal discard time based on preferences for freshness and quality, Roe, Phinney, Simons, Badiger, Bender, and Heldman (2018) find that consumers make different quality assessments when date labels are absent than when they are present. Qi and Roe (2016), almost 70% of those surveyed believe that throwing away food whose date has passed reduces the chances someone will get sick from eating the food. Wilson, Rickard, Saputo, and Ho (2017), Wilson, Miao, and Weis (2018), and Thompson, Toma, Barnes, and Revoredo-Giha (2018) find that when date labels are present, consumers often assign meaning to specific label phrases even though there is no industry standard for the meaning of these specific phrases. To decrease consumer confusion over phrases and provide more accurate information on the quality of the product, Mercier, Mondor, and Uysal (2017) suggest the use of a time-temperature indicator (TTI) label. This label communicates not only on the passage of time but also any temperature abuse to which it is exposed.

The purpose of this essay is to determine how altering milk labeling practices, in both standard and novel ways, would impact purchase and discard intentions of consumers. I use data obtained from consumers who respond to both a sensory evaluation experiment and a discrete choice experiment with milk packages featuring different levels of freshness communicated through two types of labels: a typical date label and a TTI label. From the sensory experiment, individual consumers are classified as having low, average, or high discard sensitivity based on their discard intentions. This measure is then used as an individual characteristic in the discrete choice analysis to determine how discard sensitivity

alters purchase decisions, including responsiveness to standard and novel alterations in milk date labeling. Other participant characteristics and the attributes of the milk, including the label type, freshness indicated, and price, are analyzed as well.

To the best of my knowledge, this is the first time a sensory measure of consumer sensitivity to product freshness has been used to model consumer willingness to pay. This measure partially accounts for preference heterogeneity among consumers participating in the choice experiment for packaged milk analyzed below. I argue that a robust understanding of the distribution of this preference for freshness across all consumers can yield critical insights into firm-level freshness labeling strategies and holds implications for broader policy discussions surrounding food date labeling. This is also the first study to estimate consumer willingness to pay for 'smart' freshness labels such as time-temperature indicators.

The study results show that consumers are willing to pay an additional 20 cents per carton for every additional day of freshness indicated on the package. However, the consumers in this study are not ready to pay a premium for a TTI label and, in fact, prefer the standard date label. This preference for the standard date label increases with consumer age. In addition, consumers do not value a day of freshness indicated on the TTI label any differently than on the standard date label.

The remainder of the essay is organized as follows. The next section examines the previous literature on food waste and labeling. The next section lays out the experimental design and the descriptive statistics. Then the sensitivity measure is derived from the results of the sensory experiment. This is followed by the analysis of the choice experiment

incorporating the sensitivity measure alongside milk attributes. A discussion of the results along with limitations of this study and future work concludes the essay.

2.2. Literature Review

Many authors mention consumer confusion about date labels as a significant cause of food waste (Qi & Roe, 2016; Leib, Rice, Neff, Spiker, Schklair, & Greenberg, 2016; Hall-Phillips & Shah, 2016; Newsome, Balestrini, Baum, Corby, Fisher, Goodburn and Yiannas, 2014; Gunders, 2012; Sen & Block, 2008; ReFED, 2016). However, few have attempted to investigate this confusion (Quested, 2014; Wilson et al., 2017; Wilson et al., 2018; Thompson et al., 2018).

Wilson et al. (2017), Wilson, Miao, and Weis (2018), and Thompson et al. (2018) try to pinpoint the areas of confusion by comparing consumer responses to common date label phrases including "best by," "best before," and "use by." Wilson et al. (2018) examined how consumer perceptions of product safety, quality, taste, and nutrition differed by phrase. Overall, the differences between phrases are marginal with consumers perceiving that "best by" pertains to taste while "use by" pertains to safety. These effects are not large enough to indicate a difference in how consumers would operationalize the phrases when making discard decisions. Therefore, the authors conclude that simplifying the labeling system to only two phrases could help reduce confusion about date labels but might be an insufficient policy change on its own. Furthermore, since there are no set definitions for phrases, the perceived differences between "best by" and "use by" indicate that consumer confusion about date label phrases does exist. Thompson et al. (2018) investigated the differences between "best before" and "use by" on consumers' stated

willingness to consume dairy products in Scotland where there is already a two-phrase system in place. Just as Wilson et al. (2017) and Wilson et al. (2018) concluded, Thompson et al. (2018) found mostly insignificant differences in willingness to consume between the different phrases. Thompson et al. (2018) also conclude that switching from one phrase to the other needs to be accompanied by other changes, specifically education on safety, taste, quality, freshness, and social acceptability associated with phrases, in order to significantly alter consumer behavior.

Yu & Jaenicke (2018) estimate the decrease in food waste after a law in New York City that required the sale of milk to occur within nine days of pasteurization was lifted in 2010. This resulted in the date on the package being set about six days later than before the law was lifted even though there was no change in milk quality or physical shelf life. They use a difference-in-difference model estimated with sales data from retailers and consumer panel data for milk sales before and after the law was changed. The authors found that sales fell by about 10% and that price elasticity of milk was inelastic. From this, they conclude that the change in law reduced milk waste by at least 10% if milk consumption stayed the same. If milk consumption increased, the decrease in waste could be even greater, though the authors have no direct data concerning household milk consumption or milk discard.

As shown in Roe et al. (2018), date labels on milk packages can lead to premature disposal as date labels are one of the only indications besides sensory evaluation that consumers can use to evaluate product freshness. However, in the market, these date labels are inconsistent in meaning across manufacturers. The date label is set by each manufacturer, and in states where ‘sell by’ is the common phrase on milk, is meant to guide the retailer and not the consumer. Even when the date label is consumer facing (e.g., ‘best

if used by'), there is no industry standard for the time horizon, e.g. the time of bottling to the labeled date. Manufacturers determine the time horizon based on when the product may begin to deteriorate in quality beyond a certain acceptable level, but this acceptable level can vary across manufacturers. Complicating the matter further, it is likely that this acceptable quality level varies across consumers as well. Another factor the manufacturer must consider when determining the time horizon is abuse the product could face along the supply chain and in the consumer's home. For examples, temperature abuse to refrigerated or frozen products along the cold chain can significantly increase the rate at which the product deteriorates. The manufacturer must consider the possibility of this temperature abuse into the time horizon when determining the date to be printed on the label, and these calculations can introduce further discrepancies in date labels across manufacturers. Still, because consumers use this date to make decisions that lead to premature food discard (Roe et al., 2018), the effects of the date label on consumer discard decisions need to be better understood. In particular, I am interested in two characteristics of a label: the type of information it conveys and the days remaining on the label, or 'freshness indicated.'

Typical date labels found on food packages indicate a date set by the manufacturer at the time of packaging. This date is not updated based on the conditions the package faced throughout the food supply chain or in the consumer's home. For example, milk that is stored at 6°C will decrease in quality faster than milk that is stored at the recommended 4°C. However, a typical date label is unable to communicate these deteriorated conditions. A solution to this is a time-temperature indicator (TTI) label (Taoukis and Labuza, 2003). This type of label adjusts the days remaining based on not only the passage of time but also the temperature at which the product is stored. These labels can be calibrated to different

products and are often used on pallets or large containers of food to communicate to distributors and retail managers, but not on individual packages encountered by the consumer. TTI labels on individual products provide more information on product quality to the consumer than a typical date label.

The other label characteristic of interest is the freshness time horizon indicated on the label, which has been shown in previous work to be of high value to consumers during purchase decisions (Verbeke and Ward, 2006) and to alter consumer interest in accepting price discounts (Aschemann-Witzel, 2018). With a typical date label, the freshness time horizon is the days remaining between the current date and the date on the label. As stated above, this is often a conservative measure set by the manufacturer and has no industry standard. I argue that consumers will value products that have more days remaining on the label at the time of purchase and will be more likely to consume a product that has more days remaining than a product that is close to or past the date on its label.

While the existing literature provides several key insights into the role of date labels on consumer purchasing intent, no previous work explores how consumer purchase intent responds to new approaches to communicate product freshness on food labels, nor does previous work pair consumer purchase intent responses with any measurement of consumer sensitivity to product freshness; this study fills that existing gap in literature.

2.3. Experimental Design

The experiment was held in August of 2017 with participants from central Ohio who regularly consume cow's milk and featured twelve 30-minute sessions with up to eight participants per session with all sessions conducted on the same day. Before beginning any

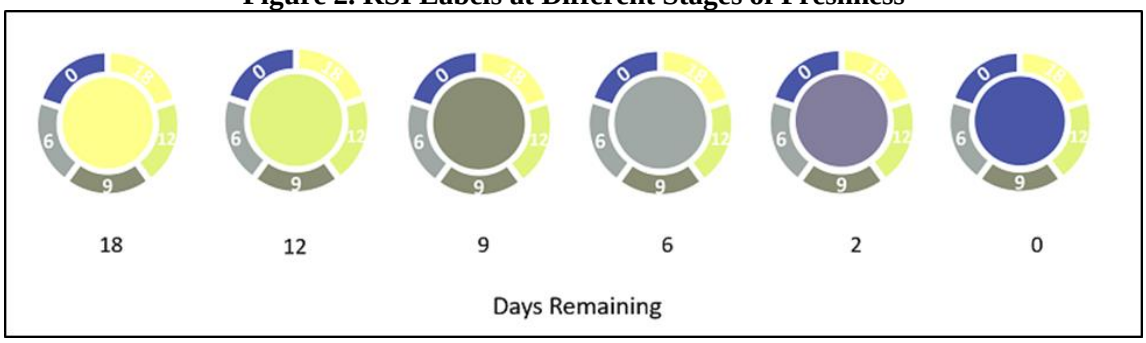
evaluations or survey questions, participants were informed they would be completing a survey about their milk consumption and that packages of milk would be presented to them during the session as part of the survey. They were instructed to not consume any of the milk. Written consent was collected from each participant. The experimental procedure was approved by the local Institutional Review Board.

Participants were placed in a booth with the survey administered on a computer in the booth. For the first stage, participants were presented with a flight of four half-gallon plastic packages of milk. The instructions were to analyze each package by visually inspecting the package and by removing the cap and smelling its contents and then indicating on the computer whether they would keep or discard the milk if the package was in their home refrigerator. Respondents were directed to sniff the inside of their wrist between each package to reset olfactory senses (i.e., this ensures one bad smell doesn't carry over to subsequent samples). The four packages of milk were commercially bottled on different days but were all stamped with a sell-by date that was 18 days post their respective dates of pasteurization, yielding packages that ranged from having three days remaining before the manufacturer's sell-by date to 22 days past the sell-by date. Two identical flights of milk were created, and then one randomly selected flight had the sell-by date, which is stamped on the package with ink, removed by the research staff. Half of the participants received the flight with the date labels first, while the other half received the flight with no labels first. The order of dates within each flight was also counterbalanced.

In the second stage of the experiment, participants were given information about current date labeling practices and a type of time-temperature indicator called the real-time

shelf-life indicator (RSI). Examples of RSI labels with differing amount of time remaining, found in Figure 1, were shown alongside an explanation of how the color changes to indicate the freshness of the milk (see Appendix A). This information was shown on the screen and given as a hardcopy handout to each participant to have for the remainder of the experiment. Following this, participants were asked six questions about the information to determine their level of understanding of the material presented.

Figure 2. RSI Labels at Different Stages of Freshness



After the quiz was completed, participants were again presented with two flights of four milk packages. Each flight consisted of two packages of milk that were in-date and two packages that were out-of-date. For each age, one package was labeled with a standard "sell by" date label while the other was labeled with an RSI label. The milk in these flights had been left out at room temperature for 24 hours the week before the experiment to simulate temperature abuse that can occur along the cold chain. The participants were unaware of this abuse for the first flight but were told the milk in the second flight had been left out overnight.

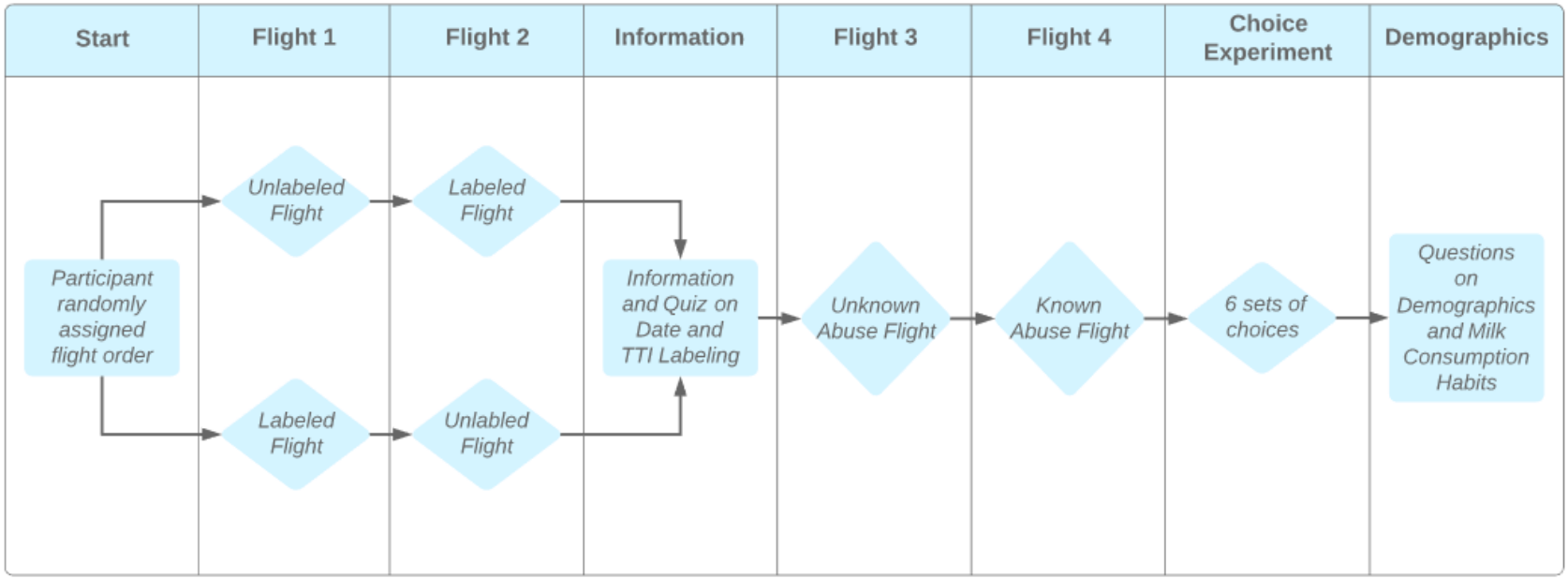
In the third stage of the survey, participants completed a choice experiment administered on the computer featuring visual representations of milk packages similar to

those used in the sensory evaluations. Half of the participants were told to imagine they were going to be out of milk (i.e., zero inventory) at home in two days while the other half were told they were currently out of milk. Then they were presented with six sets of choices and asked which milk package they would purchase at a store. A SAS macro generated a D-optimal design based on the constraints of the experiment that was used to determine the choice sets. Each choice set contained a package with a date label, a package with an RSI label, and a "neither" option. Other attributes associated with each option were price and days remaining on the label. The attributes and their corresponding levels are summarized in Table 4. A flow chart of the experiment is presented in Figure 3.

Table 4. Choice Experiment Characteristics

Attribute	Levels
Label type	Standard date label Real-time Shelf-life Indicator (RSI) label
Days remaining on label	12 days 6 days 2 days
Price	Market price 90% of market price 80% of market price 70% of market price

Figure 3. Milk Experiment Flow Chart



2.4. Data

Ninety-five participants were recruited from a large database maintained by the Sensory Evaluation Center at The Ohio State University. Participants were drawn broadly from the Columbus, Ohio community. Nine participants were dropped from the analysis of the sensory evaluation results due to failure to complete the survey or follow instructions within the survey, while another six respondents were excluded from the choice experiment analysis due to failure to demonstrate sufficient comprehension of the time-temperature indicator labels.

Table 5 summarizes the descriptive statistics of the 86 participants from the sensory evaluation. The average participant is 39 years old with the majority being female. Most participants are Caucasian (79%) with fewer being African American (9%) or Asian (9%). The average household size has 2.65 members, and 17% of the sample lives alone. As expected for the location of the testing, the majority of the sample lives in either an urban or suburban setting, while 13% live in a rural area. The average household income is $74,883, which is higher than the national average of $55,322.

Table 5. Descriptive Statistics of Participants (N = 86)

Characteristic	Description	Sample	National
Gender	Male	43%	49%
	Female	57%	51%
Age (years)	18-29	27%	22%*
	30-49	45%	33%*
	50+	28%	45%*
	Median in years	37	38
Race	White	79%	77%
	African American	9%	13%
	Asian	9%	6%
Income	Total household income	$74,883 (43,182)	$55,322**
Education	Completed 4-year college degree	73%	30%
Household size	Living alone	17%	--
	Mean	2.65 (1.21)	2.64
Residence	Rural	13%	19.3%
	Urban	42%	80.7%
	Suburban	45%	--

* Based on 2017 population estimates of adults
** 2012-2016 median household income in 2016 dollars

Because the choice experiment expresses price in terms of a percent of current market price, I estimate the absolute market price faced by each participant to develop welfare measures demarcated in dollars. I estimate the milk price for each participant based on the brand, package size, and type of milk normally purchased by the participant. The milk consumption habits of participants are displayed in Table 6. The majority of participants purchase milk from Kroger, with half-gallon packages with 2% fat content being the modal milk purchased. The average calculated price of milk normally purchased by participants was $2.44, with price ranging from $0.99 for a half-gallon of milk at Aldi to $4.99 for a half-gallon of Snowville brand milk across major retailers.

Table 6. Milk Consumption Habits of Participants

Characteristic	Description	Mean
Type purchased	Skim	13%
	1%	17%
	2%	50%
	Whole	19%
	Other	1%
Size purchased	Half Gallon	51%
	Gallon	47%
	Other	2%
Brand purchased	Kroger	52%
	Other	48%
Household consumption	Milk consumption per week	
	½ gallon or less	57%
	1 gallon	20%
	More than 1 gallon	23%
	Mean (gallons)	0.85
		(0.63)
Milk consumers	# milk consumers in household	
	1	27%
	2	45%
	3	19%
	4	6%
	5	3%
	Mean	2.14
		(0.99)
Household discard	Milk discarded in past week	
	No	60%
	Yes	35%
	Mean volume discarded (gallons)	0.28*
		(0.09)
Calculated Milk Price	Minimum	$0.99
	Maximum	$4.99
	Mean	$2.44
		(0.73)
Observations	N	86

* Number of households that discarded milk in the past week = 30.

2.5. Results

2.5.1. Participant Milk Quality Sensitivity

A distinguishing feature of this analysis is that subjects provide both preference data concerning milk via the choice experiment data as well as milk sample evaluation data from the sensory assessments conducted at the beginning of each session. The data from samples that featured a standard date label and samples where temperature abuse was not announced were used to determine each participant's discard sensitivity to milk quality.[16] Therefore, eight decisions from stage one and two decisions from stage two are used. The characteristics of the milk associated with the remaining 10 decisions used in this essay are outlined in Table 7. The quality rank was determined based on expert assessment of samples taken from each milk package used in the experiment along with the knowledge of which samples received temperature abuse as part of the experimental design (Badiger, 2018). There were four ranks of the quality of milk. Three of these, ranks associated with unabused samples, were analyzed by each participant both with and without a date label.

[16] To ensure the measure reflects responses to milk quality and not to novel information, data from packages featuring RSI labels and from any package where temperature abuse was made known to respondents was not used to develop the milk discard sensitivity measure.

Table 7. Milk Attributes

Attribute	Level
Label type	No label
	Date label
Days remaining on label	3 days
	-7 days (i.e., 7 days past label date)
	-12 days
	-22 days
Quality rank*	
(best = 1)	-7 days (abused & unabused) and -22 days (unabused)
2	Unabused with 3 days remaining on label
3	Unabused with -12 days remaining on label
(worst = 4)	Abused milk with 3 days remaining on label

* Expert assessment of quality rank did not strictly adhere to observable product characteristics due to idiosyncratic quality differences of the commercially packaged milk across production dates. For example, expert assessment placed the abused milk that was 7 days past label date in the highest quality rank ahead of the unabused milk with 3 days remaining prior to the package date.

Compared to a base group, participants classified as high (low) sensitivity are those who more (less) frequently discard milk controlling for milk quality rank and observable participant characteristics. I identify group members by analyzing the ten discard decisions discussed above. I estimate a logit model where the dependent variable is the probability of discard and the independent variables include milk attributes and participant characteristics; standard errors are clustered at the milk sample level. The results from this model are shown in Table 8.[17] This model predicts the probability of discard for each sample of milk presented to a participant. The mean of the predicted probabilities is calculated for each participant and subtracted from the actual mean of discard. Participants are classified as high or low sensitivity if this difference is at least 0.18 (one standard

[17] Other participant characteristics were not significant and dropped from the model shown. Results from the model that included other participant characteristics (e.g., living alone, household size, and milk consumption habits) alongside the results from the model in Table 8 can be found in Appendix B.

deviation) greater or less than zero, respectively. Twelve participants (14%) are classified as high sensitivity, and 13 participants (15%) are low sensitivity.

Table 8. Probability of Intended Discard

Dependent variable: Pr(Discard)		
	Logit	Marginal Effect
Milk Attributes		
Label	-0.134	-0.033
	(0.196)	
Label*Days Remaining	-0.053***	-0.013
	(0.016)	
Quality Rank	1.057***	0.257
	(0.291)	
Participant Characteristics		
Age	0.107***	0.026
	(0.027)	
Age^2	-0.001***	0.000
	(0.000)	
ln(income)	-0.229***	-0.056
	(0.061)	
Female	0.056	0.014
	(0.096)	
College degree	0.119	0.029
	(0.195)	
Rural	0.213	0.052
	(0.272)	
Suburban	0.219*	0.053
	(0.132)	
Constant	-1.520	
	(1.043)	
Psuedo R^2	0.069	
Observations	880	

*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the milk sample level.
Marginal effect at the mean reported.

Intuitively, the high sensitivity participants are those consumers who are more conservative in their discard decisions. They tend toward discarding milk more than other participants, all else equal. The low sensitivity participants are the opposite. These consumers are more likely to keep milk, all else equal. This measure is included as an individual characteristic in the choice experiment analysis to account for consumer preference heterogeneity that would normally be swept into the unobservable term.

2.5.2. Purchase Intent

In the choice experiment, participants are presented two milk purchase options in addition to an option to forgo immediate purchase. Assuming a standard random utility model framework, participants choose the alternate that maximizes their utility. Specifically, following the model from Train & Weeks (2005), the utility that consumer *i* receives from alternate *j* in the choice set *t* is

$$U_{ijt} = -\alpha_i p_{ijt} + \beta_i' x_{ijt} + \varepsilon_{ijt} \quad (1)$$

where p_{ijt} and x_{ijt} represent price and non-price attributes, respectively. The attribute parameters, α_i and β_i, are individual specific; and ε_{ijt} is unknown and independent and identically distributed. The variance of ε varies over consumers if consumers differ in the unobservable factors that affect their decisions, thus inserting a degree of randomness in analysis since these factors are unobservable to all but the consumer. The consumer chooses alternate *j* over *k* when the utility received from alternate *j* is greater than the utility received from alternate *k*. Therefore, the probability of consumer *i* choosing alternate *j* in a given choice set is

$$Prob(U_{ij} > U_{ik}), j \neq k \quad (2)$$

To calculate these probabilities from the choice experiment, a mixed logit model is used. This model is chosen as it accommodates preference heterogeneity by allowing for varying individual coefficients and is not restricted by the independence of irrelevant alternatives property meaning that substitution patterns are not restricted (McFadden & Train, 2000). The specified model assumes a fixed distribution of the parameter for price with all other attribute parameters assumed to be independently and normally distributed and allowed to vary across individuals.

The purchase decision is the dependent variable, and price is the calculated price discussed in the previous section. An individual's willingness to pay (WTP) for each attribute is calculated from the estimates $\beta_{attribute}$ and α obtained in equation (1) and is represented by the ratio

$$WTP_{attribute} = \frac{\beta_{attribute}}{\alpha} \quad (3)$$

Confidence intervals for estimated WTP are constructed using the delta method implemented via STATA 13.1 (Hole, 2007).

The results of the mixed logit models from the 80 qualifying respondents are shown in Table 9 along with estimated willingness to pay.[18] Three of the four attributes — price, days remaining, and the option to purchase neither — have a significant effect on the purchase decision.[19]

The marginal effect for price[20] of -3.565 is highly statistically significant. Using this estimate, the willingness to pay can be calculated for the other attributes. The consumer with a standard level of discard sensitivity is willing to pay \$0.20 per carton for each additional day remaining on the label, with the 95% confidence interval ranging from \$0.15 to \$0.25 for each additional day. Participants identified as having low discard sensitivity valued the days remaining by \$0.06 per carton less per day on average than consumers with a standard level of discard sensitivity, while participants with high discard sensitivity valued days remaining by \$0.03 per carton more per day on average than consumers with

[18] Results featuring all participants are included in Appendix C for reference; coefficient estimates and willingness to pay values are largely unaffected.
[19] The estimated coefficients from participants who were informed that they were currently out of milk were not statistically different from those who were told they would be out of milk in two days. To simplify exposition, this factor is not included in the model.
[20] Price is assumed to have a linear and normally distributed coefficient.

a standard level of discard sensitivity, though the confidence interval for high sensitivity participants includes both individuals who valued additional days both more and less than the average sensitivity (base) group.

The RSI label attribute marginal effect is negative and significant for the base oldest age category (age 51 and older) with a willingness to pay of -$0.56 and a confidence interval of -$0.96 to -$0.17. The standard deviation coefficient for the RSI label attribute is not significantly different than zero, implying limited heterogeneity in the disutility caused by this label once participant age (discussed below) is controlled. This means that consumers would be willing to pay $0.56 for the date label instead of the RSI label. Age is a clear mediating factor in preferences for the RSI labeling technology. Specifically, the interaction effect between the RSI label and those participants age 29 and younger is positive and significant, with willingness to pay for the RSI label increasing on average by $0.39 for participants in this age group compared to the base group of participants more than 50 years of age. The average participants aged 30 to 49 had a willingness to pay that was $0.27 more than the oldest group, though the confidence interval for this middle age group featured some participants with willingness to pay that was higher than the oldest group.

Table 9. Mixed Logit Model of Purchase Intent and Marginal Willingness to Pay

Variable	Marginal Effect	SD	Willingness to Pay ($/carton)
Prob(Purchase=1)			
Price ($/gallon)	-3.565*** (0.646)		
Label Days Remaining	0.726*** (0.107)	0.336*** (0.072)	-$0.20 [0.15, 0.25]
Days Remaining x Low Discard Sensitivity	-0.224* (0.128)		-$0.06 [-0.13, -0.01]
Days Remaining x High Discard Sensitivity	0.118 (0.192)		$0.03 [-0.07, 0.13]
RSI Label	-2.013*** (0.638)	-0.040 (0.424)	-$0.56 [-0.96, -0.17]
RSI x Young Age Category	1.380** (0.459)		$0.39 [0.05, 0.72]
RSI x Middle Age Category	0.961* (0.525)		$0.27 [-0.03, 0.57]
Neither Package	-9.473*** (1.932)	5.073*** (1.160)	-$2.66 [-3.32, -1.98]
Log Likelihood	-219.718		
AIC	463.435		
BIC	526.704		
Observations	480		480

*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.
Marginal effects are reported in the first column along with standard errors in parentheses.
95% confidence intervals reported in square brackets.

Finally, the marginal effect of purchasing neither package was highly negative and significant, meaning that consumers were willing to pay $2.66 on average to purchase one of the two options rather than "leaving" the hypothetical shopping scenario without purchasing either package of milk. Versions of the model were also estimated that included an interaction between the RSI label and the number of days remaining on the label;

however, neither the marginal effect nor the standard deviation parameter was statistically significant and hence, these were omitted from the model for the sake of simplicity. This term was expected to be positive and significant since the days remaining indicated on the RSI label is meant to be more accurate than the days remaining on the date label, which may suggest that the full benefits of this advanced labeling technology were not fully understood by participants or not fully trusted by the participants if they were understood.

2.6. Discussion

The results indicate significant preference heterogeneity among milk consumers with respect to the number of days remaining. This is because the participant discard sensitivity interaction term and two coefficients of standard deviation were statistically significant in the mixed logit estimation.

Manufacturers and retailers who are aware of these different types of consumers and how each type values product attributes such as days remaining might be better able to implement strategies to both reduce waste and maximize store profits. Any adjustment made will have differing effects on different types of consumers. The first part of the experiment allowed categorization of participants by discard behavior, which I call their discard sensitivity. This classification controls for expert assessment of milk quality and observable attributes such as bottling date and captures previously unobserved sensory-based responses that translate to discard intentions. This result provides evidence for the first time on the relationship between discard tendencies and purchase decisions of consumers.

This discard sensitivity partially explains the preference heterogeneity observed across consumers. Consumers with low discard sensitivity care less about the days remaining on the label, which means that compared to those with average or high discard sensitivity, these consumers are willing to pay less for packages where more days remaining on the label and are likely to respond less to price discounts for milk that is nearer to its package date. These consumers are also less inclined to waste product, so food waste reduction strategies do not need to prioritize targeting this group.

The high sensitivity group is the opposite. This group should be targeted as these consumers tend toward discarding product. While there is no significant difference in their valuation of days remaining compared to average sensitivity consumers, they do have a positive, insignificant coefficient on the days remaining with a majority of such participants holding a WTP that is higher than participants with medium discard sensitivity. This indicates that they may be the group most sensitive to the days remaining on the label. Therefore, an increase in the freshness indicated on a label could increase purchases made by this group, while they might only respond to milk with a limited freshness horizon if it is steeply discounted. All else equal, an increase in purchases within this group would lead to an increase in food waste because of their "quick trigger" during the discard intention (sensory) experiment.

One simple strategic adjustment retailers could consider to decrease food waste is to increase the freshness indicated on the package. For example, Yu and Jaenicke (2018) estimated that when New York retailers increased package dates from nine days post bottling to an average of 15 days, household-level waste declined by about 10%.

For the bottler producing the milk used in this experiment, the time horizon from the day of bottling to the date printed on the package is 18 days. Based on the results, if this time horizon increased to 20 days, the retailer could charge up to $0.40 more per package as this would translate to two additional days of freshness indicated on the package. However, such a move increases the risk of consumers discarding the milk before it is consumed in the household due to deteriorating sensory characteristics. In this event, the reputation of the retailer could be damaged and hurt future sales.

How much additional discard would occur is an empirical question that would require additional laboratory or market level experiments. Understanding these impacts are critical, however, as milk is considered a loss leader, i.e., a good that is sold at lower prices to bring consumers into the store thus leading to sale of other goods that are profitable. Any damage in reputation a retailer experiences from milk spoilage before the date on the package would not only lead to a loss in milk sales but also in the sale of other goods purchased by consumers who were initially "led" into the store by the price or quality of the milk.

Another strategy milk producers and retailers could wield to reduce food waste while minimizing the risk of milk going bad before the date on the package is the adoption of the RSI label. The RSI label is an advancement especially suited to milk. Quality degradation of milk throughout the cold chain is highly dependent on storage temperature. The RSI label can accurately capture this degradation and, at a cost of pennies per package, is not cost-prohibitive. However, based on the results from this particular sample and protocol, switching to an RSI label would decrease milk sales. Young and middle age consumers do not show a preference for one label type over the other, but older consumers

show a clear preference for the date label. Therefore, stores that switch to the RSI label would need to alter demand or lower the price. Furthermore, consumer demand for the RSI label could be altered with more education about the label or with increased experience with the RSI label.

Another strategy available to retailers and producers is to add the RSI label in addition to the date label. In this experiment, only one label was applied to each package. This created a trade-off in information the consumer received from the package. With the date label, consumers had some sense of when the milk was bottled based on the date printed on the package. With an RSI-labeled package of milk with no shelf life remaining, the consumer is unaware of how far "expired" the milk is. Even before the RSI label expires, the consumer is unaware if the milk is near the end of its shelf life because of temperature abuse or because of general age. Including both a date label and an RSI label on the package could increase consumer knowledge of the quality of the product without decreasing sales.

If producers switched to the RSI label solely, not only could sales decrease, but food waste could also increase. This would occur if there were no change in time horizon used to calibrate the RSI label. Producers purposely use conservative time horizons on date-labeled milk to decrease the risk of a consumer discovering the milk has spoiled before its printed date. Since the producer only has one opportunity to set this printed date at the time of bottling, they must factor in temperature abuse that can (and often does) occur along the cold chain including temperature abuse in the consumer's home.

The RSI label responds in real-time to temperature abuse. If this time horizon is not adjusted when switching to the RSI label, then packages of milk that are not temperature

abused will show no days remaining on the label even when they are still fresh. Temperature abuse will be accounted for through not only the conservative time horizon, but also through the increased progression of the RSI label through its life cycle. This overlap would increase the amount of food discarded when it is still fresh thus causing an increase in food waste. Therefore, if a retailer were to adopt RSI labels and wished to reduce the amount of milk wasted by consumers at home, adjustments would need to be made to the time horizon.

The time horizon used for an RSI label would need to be longer than the time horizon used for a date label for this switch to decrease food waste. The time horizon chosen does not have to account for temperature abuse that can occur along the cold chain. If a package is temperature abused, the RSI label will reflect that by displaying fewer days remaining. If a package is not temperature abused, the RSI label will display the days remaining that correspond with the time horizon chosen. Since the time horizon represented by the RSI label does not need to account for temperature abuse, it can reflect the true shelf life of a properly stored package. Using a longer time horizon for an RSI-labeled product than a date-labeled product could affect sales and waste in different ways. It is unlikely that the RSI label would increase the amount of milk discarded at the household level due to the label capturing temperature abuse since manufacturers currently use conservative time horizons. Milk waste at the household level could decrease as consumers would be less likely to throw out fresh product based on the label. This decrease in waste would translate into lower sales for retailers.

The relationship between food waste and sales is a concern for any retailer-driven food waste reduction strategy. Specifically, the feasibility of such a strategy is

questionable. If consumers waste less food, then the demand for food shifts to the left. This will decrease the total sales for the retailer. Grocery stores that act as both the producer and the retailer are unlikely to implement a strategy that leads to a loss of sales unless there are benefits through a better reputation or a higher margin per unit sold. Currently, retailers such as Kroger and Walmart are embracing campaigns to make consumers aware of the retailers' efforts to curb food waste. These campaigns include information about the problem of food waste in general as much of the public does not recognize its harmful effects. To this end, the retailers are taking on the responsibility of creating consumer demand for food waste reduction as well. If adequate demand is generated, then measures to reduce consumer food waste could prove profitable for producers.

2.7. Conclusion

Wasted food generates significant financial, environmental, and societal costs. Many food waste reduction strategies have been proposed, but their impact needs to be understood before they are implemented. This study suggests that increasing the freshness time horizon on date labels would decrease food waste in the home. It also indicates that consumers differ in their preferences for product freshness. These differences not only affect purchase decisions but also discard decisions. For this reason, producers implementing food waste reduction strategies need to consider the impact on each type of consumer to minimize food waste reduction while maximizing profits. This preference heterogeneity also needs to be accounted for in future studies analyzing policy interventions to accurately assess the effect of those interventions on household and aggregate level food waste.

This study also suggests that consumers, especially older consumers, are not ready for smart labels such as the TTI label. This novel labeling process would need to be introduced into the market with caution. Educational campaigns on this labeling style targeting the older demographic or adding a novel label alongside a traditional label to alleviate information trade-off could improve consumer acceptance, but future work is needed to determine the effectiveness of such tactics. In addition, further work on other cost-effective packaging solutions that provide accurate information on the freshness of the product to the consumer is required to identify packaging that might be more readily accepted by consumers.

Chapter 3.

Date Labeling Environment: An Analysis of Phrase-date Combinations on Consumer Discard Behavior

3.1. Introduction

A consumer's decision whether to discard a food product is based on preferences for freshness and quality. Further, there is evidence that consumers make different quality assessments when date labels are absent than when they are present (Roe et al., 2018). When date labels are present, consumers often assign meaning to specific phrases they see even though there is no industry standard for the meaning of any phrase (Wilson et al., 2018; Thompson et al., 2018). Proposals to combat the effect of this confusion range from standardizing the date labeling system to moving the location of certain labels to less visible parts of the package to simple educational campaigns about date labels or food waste (Leib et. al, 2016; Newsome et. al, 2014). Many reports call for the standardization of labels or a two-phrase date labeling system and suggest that this is the most cost-effective way to decrease food waste (Lipinski, Clowes, Goodwin, and Hanson, 2017; Leib et. al, 2016; Newsome et. al, 2014; Gunders, 2012).

In 2017, the Food Marketing Institute (FMI) along with the Grocery Manufacturers Association (GMA) introduced a voluntary initiative to adopt a two-phrase date labeling

system in the United States. For foods that have higher safety risks associated with product age, FMI/GMA proposed "Use by" as a phrase that indicates a product safety measure. For foods that may deteriorate in quality but do not increase in safety risks with age, the organizations proposed "Best if used by" to indicate a product quality measure (GMA & FMI, 2017). While this proposal has been praised for its attempt to clear up consumer confusion over date labels, much about its impact is unknown. In May 2019, the FDA sent a letter to the food industry supporting the use of the phrase "Best if used by" for indicating quality of food packages and encouraged its adoption; however, the FDA explicitly did not recommend the phrase "Use by" as per the FMI/GMA proposal due to "safety reasons at this time" (Yiannas, 2019). It appears that the impact of switching to a two-phrase system on consumer food discard behavior needs to be better understood before the FDA will fully endorse it. While the two-phrase system has not been directly studied, some studies have attempted to decipher what differences consumers interpret from different phrases in the current unstandardized environment.

Wilson et al. (2018) and Thompson et al. (2018) try to pinpoint the areas of confusion for consumers by comparing "best by" or "best before" to "use by." Wilson et al. (2018) examined how intensely consumers perceived the phrases to be related to safety, quality, taste, and nutrition. Overall, the differences between phrases are marginal with consumers perceiving that "best by" pertains to taste while "use by" pertains to safety. These effects are not large enough to indicate a difference in the actual meaning of the phrase. Therefore, the authors conclude that simplifying the labeling system to only two phrases could help clear up some confusion about date labels but would not be a sufficient policy change on its own. Furthermore, since there is no set definition for phrases currently,

the perceived differences between “best by” and “use by” indicate that consumer confusion about date label phrases does exist.

Thompson et al. (2018) investigated the differences between “best before” and “use by” on consumers’ willingness to consume dairy products in Scotland where there is already a two-date system in place. Just as Wilson et al. (2018) discovered, Thompson et al. (2018) found mostly insignificant differences in willingness to consume based on the different phrases. Thompson et al. (2018) also conclude that switching from one phrase to the other needs to be accompanied by other changes, specifically education on safety, taste, quality, freshness, and social acceptability associated with phrases, to significantly alter consumer behavior.

In this essay, I analyze the effect of the date label on consumer discard intention in the time period surrounding that of the date on the label by manipulating the phrase and date on the label of five different products in a sensory laboratory experiment. Through this variation in the phrase-date combination, I determine that the effect of the label on intended discard is not impacted by the phrase on the label but rather is driven by the date on the package. Information regarding the meaning of label phrases and how the date is chosen does not affect consumers’ intended discard behavior.

The rest of the essay is organized as follows. Section 3.2 discusses the possible mechanisms through which date labeling affects consumer discard behavior. Section 3.3 lays out the experimental design for this essay followed by the data and summary statistics in Section 3.4 and empirical model in Section 3.5. Section 3.6 provides the results with Section 3.7 concluding the findings of this essay including policy implications, discussing limitations, and suggesting future research.

3.2. Mechanism

The phrase-date combination on labels could have several effects on consumer behavior both at the point of sale and within the home. At the point of sale, a product with a phrase-date combination that the consumer perceives as indicating a fresher product is demanded more than a combination that indicates an older product. Chapter 2 of this dissertation found that while consumers value an extra day of freshness, some place a higher value on it than others depending on their discard behavior in the home. This essay focuses on this discard behavior in the home.

Several consumer behaviors in the home may be impacted by the phrase-date combination on the label. First, consumers could use the label to determine when they will consume the product throughout its shelf life. If three boxes of the same cereal are in their pantry, a consumer may check the label to determine which one is the oldest, and then first open that box of cereal to consume. Second, and closely related, consumers could use the label to prioritize the use of a particular ingredient as the date on the label draws near. A consumer may know that their milk is getting old and choose to use it in a recipe or on their newly opened cereal. Finally, consumers may use the label to aid in the decision to discard any remaining product due to quality or safety issues. If the label indicates the product is older, a consumer may immediately discard the product or carefully inspect the product before deciding whether to consume, keep, or discard it. The latter two behaviors can trigger consumers to purchase more product to bring into the home with the intention of consuming those products. While the first two behaviors revolve around meal planning which could lead to future food waste, the final behavior is the actual act of creating food waste through discarding food and is the main behavior this essay investigates.

Consumers can use their senses, information on the package, or both to make a discard decision about a product. Roe et al. (2018) found that when consumers were only able to use their senses to evaluate milk packages they were less likely to keep an in-date product and less likely to discard a past-date product than when date information is available on the package. This indicates that consumers are assigning some meaning to the phrase-date combination on the label and using that information in their discard decisions. However, the intended meaning of phrase-date combinations is inconsistent at best, varying from manufacturer to manufacturer, and misleading at worst, such as when a phrase meant for retailers (e.g. “sell by”) is placed on a central portion of the product package in clear sight of potential consumers.

A lack of regulation or standardization surrounding date labeling combined with poor consumer understanding and interpretation of the date label leads to safe, high-quality food being discarded (Thompson et al., 2018; Newsome et al., 2014). Consumers tend to assume that the phrase-date combination on a package has a specific meaning. Since manufacturers are free to put any phrase-date combination that comports with relevant state-level regulations on the package, the certainty of the label’s meaning leads consumers to interpret that the close-to-date or past-date products are unsafe or of a lower quality and, as a result, discard it.

3.3. Experimental Design

Participants were recruited from a large database maintained by the Sensory Evaluation Center at The Ohio State University and drawn broadly from the Columbus, Ohio community. Participants were required to either consume or prepare the products

being tested on a regular basis, have no sensory deficits, and be 18 years of age or older. Testing of purchase behavior was performed over the same two weeks, so an additional requirement was that participants could not have participated in the purchase version of the study. Participants were only allowed to be tested over one set of products as well to ensure they were not exposed to two treatment groups. Overall, 364 people met the requirements and participated in the discard behavior portion of the experiment; however, 17 participants were dropped from the analysis due to failure to complete the survey or follow instructions within the survey.

The experiment was conducted over two weeks in August of 2018. There were four days featuring eight to ten 30-minute sessions per day. Up to 10 participants were tested per session. Before beginning any evaluations or survey questions, participants were informed they would be completing a survey about their consumption of specific products and that packages of these products would be presented to them during the session as part of the survey. They were informed that they could open, smell, and visually inspect each package except for chicken, but they were not to consume any product. The packages of chicken could be smelled and visually inspected, but not opened or consumed. Written consent was collected from each participant before they entered the sensory laboratory. The experimental procedure was approved by the local Institutional Review Board.

Participants were placed in a booth with the survey administered on a computer in the booth. The participants were presented with a flight of three packages of the same product. The survey then prompted the participant to analyze each package and indicate on the computer whether they would keep or discard the product if the package was in their home refrigerator.

Most of the products[21] were identical and packaged on the same day with the only difference being the phrase-date combination printed on the package. All the products were "in-date" in terms of the original date printed on the package label by the manufacturer. This label was manipulated by the research staff to show three different "ages" of product with differing accompanying phrases. Each flight had products of the three labeled dates representing the different product ages. The phrase used within each flight was the same for each participant but differed between groups of participants. These phrases by product and group is presented in Table 10.

Three sets of flight characteristics were created: one with no phrase or date, one with the standardized phrase for that product, and one with the most commonly used phrase that was not the standardized phrase for that product – referred to as the "status quo" label for the remainder of this essay. Two additional treatment groups were created from providing information about date labeling to the participants. Information was only given to those who saw packages with labels.

In total, six products were tested with each participant seeing one set of three products. The first two days of testing featured chicken, lettuce, and bread while the final two days featured frozen chicken, milk, and cereal. These products were chosen based on discussion with industry partners and represent different categories of packaged foods found grocery stores: fresh meat (chicken), packaged produce (lettuce), bakery (bread), a

[21] Breakfast cereal was the only product that was not identical in actual age. The actual age varied by 2 weeks with all the packages being at least 6 months from the manufacturer's date label (i.e. the product was "fresh"). The cereal was used in the experiment while it was still early in its long, relatively stable shelf life and remained unopened until two days prior to the day it was used for testing; therefore, I assume the cereal is all of the same quality.

frozen item (frozen chicken), dairy (milk), and a center-store/dry-goods item (cereal). The order of dates within each flight was counterbalanced, and the order the products were presented was reversed between the first and second day that set of products was featured. For this essay, only chicken, lettuce, bread, milk, and cereal will be examined.[22] The days remaining are calculated as roughly 10% of the shelf life pre-date, 10% of the shelf life post-date, and 30% of the shelf life post-date. Table 10 summarizes the phrase-date combinations used by product and the information received for each group of participants.

Table 10. Group Characteristics for Phrase-Date Combinations and Information

		No Label	Status Quo Label		Standardized Label	
	Days Remaining	Group 1	Group 2	Group 3	Group 4	Group 5
Information		No	No	Yes[1]	No	Yes[2]
Chicken	+1, -1, -3	-	Sell by	Sell by	Use by	Use by
Lettuce	+1, -1, -3	-	Sell by	Sell by	Best if used by	Best if used by
Bread	+1, -1, -3	-	Best before	Best before	Best if used by	Best if used by
Milk	+2, -1, -7	-	Sell by	Sell by	Best if used by	Best if used by
Cereal	+20, -10, -70	-	Best before	Best before	Best if used by	Best if used by

[1]Information on the unstandardized, "status quo" date labeling system

[2]Information on standardized date labeling in a two-phrase labeling system

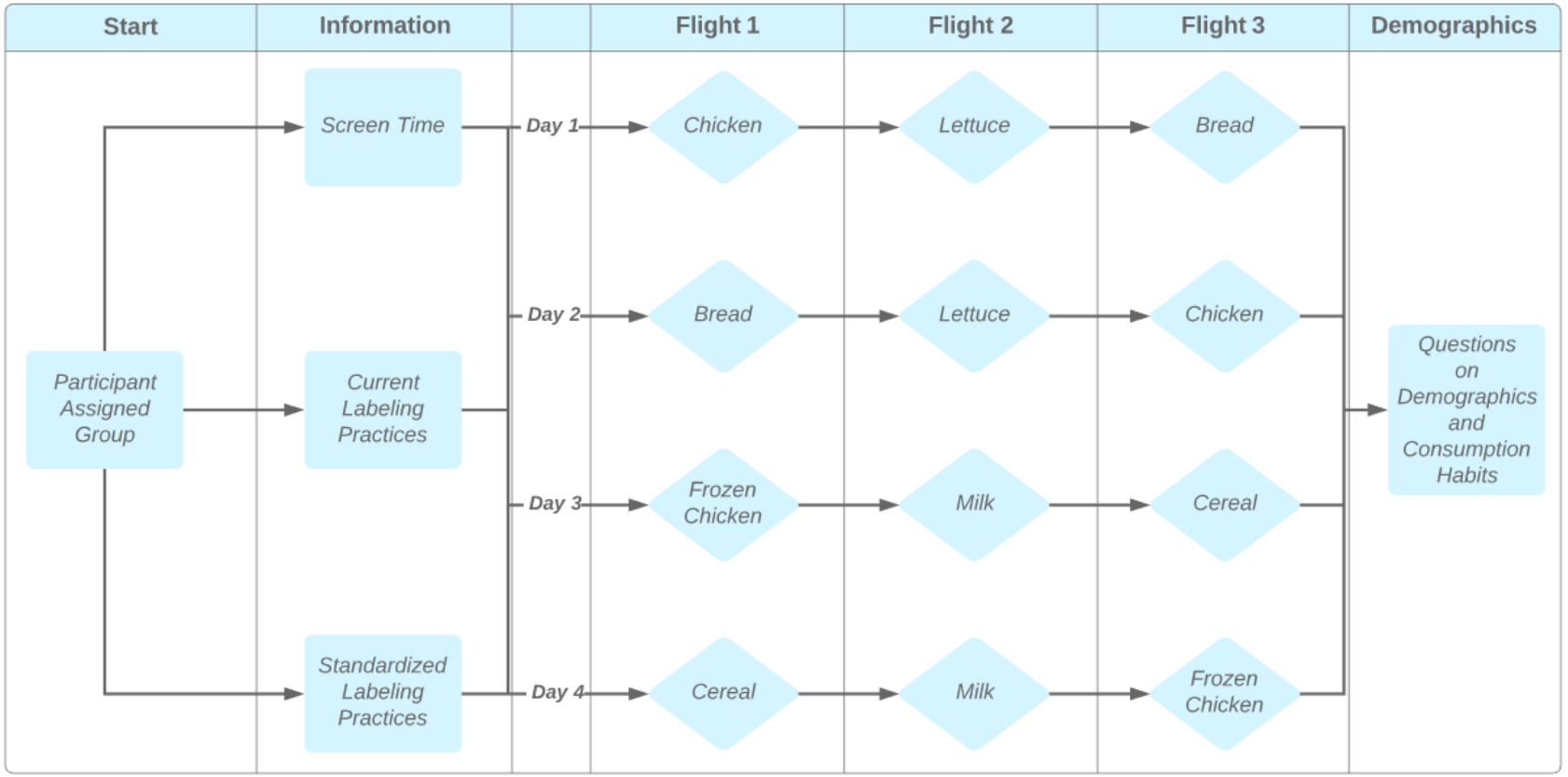
Information was provided at the beginning of the survey. Group 3 received information regarding the current date labeling environment and Group 5 received information about the two-phrase, standardized labeling system. The information given to each of these two groups can be found in Appendix D. The other three groups were given information on screen time for children, a topic unrelated to the topic of interest. This ensured that everyone had a similar experience progressing through the beginning of the

[22] Frozen chicken is a distinct category of food tested in this experiment. It was packaged and labeled in a different way than the five other products. Participants were told that the chicken opened, partially used, and then placed in a freezer prior to the date on the package. Unlike the other products tested, all labels on the frozen chicken were past-date.

survey. All the participants were quizzed with six questions about the information they received to check their understanding and comprehension of the material. A flow chart of the experiment is shown in Phrase-Date Experiment Flow ChartFigure 4.

The product label was manipulated differently for each product based on its original packaging. The chicken was packaged by a local Kroger store, and the label was left off the package. Research staff then affixed the appropriate label sticker to each package. The original labels on both packaged lettuce and milk were removed with acetone then reprinted with a handheld printer. A local Kroger store donated bread along with extra unlabeled bread bags. The handheld printer was used to label these bags, and then the bread was moved into the new bag. Unlabeled cereal boxes could not be obtained from a manufacturer, so the research team designed and had printed a cereal box that reflects industry standard branding practices. The handheld printer was again used to print the appropriate labels on the new packages. For all products, the font and size of the text was chosen to most closely resemble the original label style. Pictures of product packaging can be found in Appendix E. Approximately one-third of product was removed from each package to simulate an opened package the participant might encounter in their home.

Figure 4. Phrase-Date Experiment Flow Chart



3.4. Data

Table 11 summarizes the descriptive statistics of the 347 participants used in the analysis of the sensory evaluation results. National statistics are in the last column for comparison. It is important to note that these national statistics are based on the population as a whole rather than the population that consumes or prepares the products in the evaluation. This could partially explain differences between the sample and national statistics. The sample consists of about 65% females and most of the participants are over the age of 50 (65%). In line with the national population, the majority of participants are Caucasian (75%) with a lower percentage of black (11%) and Asian (9%) participants. The percentage of participants with a four-year college degree is higher than the national average which is most likely due to the sample being drawn from the area around The Ohio State University. In addition, the median income for the sample is about $30,000 higher than the national median.

Table 11. Summary Statistics (N =347)

Characteristic	Description	Sample	National
Gender	Female	64.7%	49.0%
	Male	35.3%	51.0%
Age	18-29	13%	22%*
	30-49	22%	33%*
	50+	65%	45%*
	Median age	41	38
Income	Median household income	$85,000	$55,322**
Education	Completed 4-year degree	76%	30%
Household size	Mean household size	2.8	2.6
	Living alone	13.2%	--
Kids	Mean number of children (<18 years) in household	0.46	2.6
	Percentage of households with children	26.9%	--
Race	White	75%	77%
	Black	11%	13%
	Asian	9%	6%

*Based on 2017 population estimates of adults
**2012-2016 median household income in 2016 dollars

3.5. Empirical Model

To determine if the label impacts consumer discard intention, I estimate a logit model of all the products combined and then separately for each product. This model is specified as:

$$\pi_{ij} = \beta_0 + \beta_1 Label_j + \beta_2 Label_j * Age_j + \beta_3 Label_j * Info_i + \beta_4 Label_j * Info_i * Age_j + \gamma \boldsymbol{X}_i + \varepsilon_{ij} \quad (1)$$

where *j* is the package being examined by individual *i*; π is the probability of the decision made by individual *i* on package *j* to keep or discard the package; *Label* represents the dummy variables for the type of label on the product package which is either the status quo label or the standardized label; *Label*Age* is interaction between the label dummies and the dummies for product age, *young*, *medium*, and *old* or *pre-date* and *post-date*; *Info* is if the participant received information on date labeling at the start of the survey;

*Label*Info*Age* is the interaction between label type, product age, and if the person received information on date labeling at the beginning of the survey; and $\boldsymbol{X}$ is a vector of participant characteristics including age, gender, kids living in the household, education, and income.

For this essay, frozen chicken is dropped from all analyses so the products remaining are fresh chicken, lettuce, bread, cereal, and milk. There are nine observations recorded for each participant given the fresh chicken, lettuce, and bread set of products and six observations for each participant given the frozen chicken, milk, and cereal set of products. All models in this essay have standard errors clustered at the session level.[23]

In the second phase, I estimate the same model but by product. This allows for analysis of effects of label, product age indicated on the label, and information on each specific product as the estimates can now vary across the subsamples of products.

In the final phase, I use the same model in the second phase but with a different dependent variable. I attempt to identify socially optimal consumer behavior of managing food. Externalities associated with food production and distribution are not accounted for in the market price of the food. These externalities include, but are not limited to, soil erosion, decreasing water tables, and greenhouse gas emissions from agricultural production, air pollution from distribution along the supply chain, and methane released from food decomposing in landfills. Reducing food waste is a public policy goal in the United States (U.S. Department of Agriculture, 2016). One complication to meeting goals to reduce food waste is that there is no set definition of *food waste*. The results of policy

[23] There were ten sessions on the first day of testing and eight sessions on the second day of testing for each of the two weeks for a total of 36 sessions. Sessions included two booths of each of the groups (1-5) of participants.

changes can vary based on the definition or measurement of food waste that is used. In the previous sections, I examined the effect of information and labels on discard intention; however, while the strictest of definition of food waste would equate discard with waste, it is generally agreed upon that some food being discarded is socially acceptable and is not waste. For example, one definition of food waste does not consider stems, pits, and other inedible parts of food to be waste. The externalities associated with food waste are not accounted for in the price of the food, but these inedible parts of the food were never meant for human consumption.

In addition to this, it is sometimes socially optimal for edible parts of food to be wasted. The elderly and those with compromised immune systems have greater risks associated with consuming older foods. The medical costs associated with these risks could outweigh the benefits of the food not being discarded. Avoiding the externalities of the food being discarded, such as the methane released from the food decomposing, would need to be weighed against the expected health costs of consuming the food for the compromised group of people. Weighing the public policy objective to reduce food waste with the objective to improve public health is important; however, some policies may improve both. Less consumer confusions around date labels could decrease food waste and improve public health simultaneously.

In this section, I analyze the socially optimal management of food by defining a new dependent variable *mismanage*. A product is mismanaged any time fresh food is discarded or unsafe food is kept. It is this second category of unsafe food that is kept that alters the *mismanage* outcome variable from the *discard* outcome variable used so far in the analysis. *Mismanage* is defined here to capture the socially optimal consumer behavior.

With the standardized label, the date on the package is the date that the social cost of negative health effects outweighs the social benefit of consumption. All the products in this experiment are fresh so the only packages that should be discarded are those with past-date standardized safety labels. Of the five products tested, chicken is the sole product that has a standardized safety label. Therefore, this analysis is limited to fresh chicken as the results for other products would not differ from those above where *discard* is the dependent variable since *mismanage* is equal to *discard* for those products.

3.6. Results

3.6.1. Intended Discard

Participants were asked to examine each package of product presented to them and decide what they would do with that package if it were in their home. The options they could choose from were "keep," "discard," or "unsure." If "keep" or "unsure" was chosen, then participants were asked a follow up question about the amount of time they estimated they would keep the product before discarding it. The outcome variable *discard* used in this analysis is created from the answers to these two questions. *Discard* is coded as 1 if the participant answered "discard." If the participant answered "keep" or "unsure" and the participant estimated that they would keep the product for an additional nonzero number of days, then the *discard* is coded as 0. If zero was entered for the additional number of days that the product would be kept, then *discard* is coded as 1. There were 20 instances where the participant did not choose "discard" but were coded as choosing "discard" through their answer to the follow-up question.

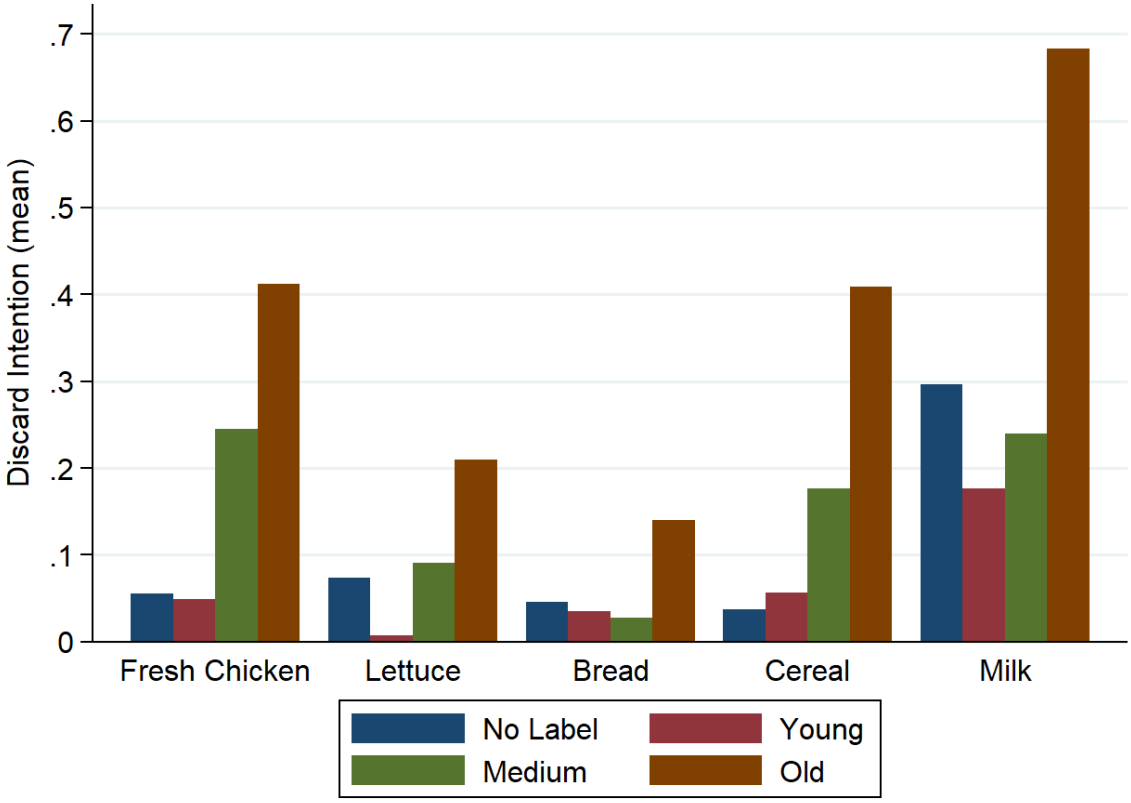
Figure 5 shows the mean discard rate by product for each of four groups of packages: those with no labels and those with labels broken down by the product age indicated by the label: young, medium, and old. When deciding on a package with no label, consumers could only rely on their senses while consumers who received packages with a label have the additional information they interpret from the label. All the products were within the original date printed by the manufacturer so none of them needed to be discarded due to quality or safety concerns. When no label is present, milk has the highest rate of discard at 31.7%, followed by lettuce at 7.4%, fresh chicken at 5.7%, bread at 4.6%, and finally, breakfast cereal at 3.7%.

The labeled products are divided by product age. When a product has a pre-date, or young, label, the rate of discard intention falls from the rate of the unlabeled group for all products except cereal which had a slight increase to 5.6%. Cereal is the most shelf-stable of the products used in the experiment and one that participants consider to be "safe." Participants indicated tasting the cereal was very important to them when making a discard decision; in this experiment, they were restricted from doing so. This, combined with low actual and perceived health risks of consuming old cereal, could account for the low discard intention rate among those who could only use their senses. The rate of discard for milk fell but is still the highest of all products at 17.6%. Chicken, lettuce, and bread were all lower than their unlabeled counterparts at 4.9%, 0.7%, and 3.5%, respectively. Lettuce had a particularly low rate with only one participant choosing to discard a young-age lettuce package.

Packages with post-date labels overall had a higher discard rate for all products. In Figure 5, the young sample includes only those that are about 10% of the shelf life pre-date

while the post-date labels include the medium- and old-age packages that had dates that were about 10% and 30% of the shelf life past-date. Medium-age bread has an intended discard rate slightly below that of the young-age bread before showing a sharp increase in discard for old-age packages. All the other products exhibit an increase in discard between the young and medium ages and then again between the medium and old ages. Among the old-age sample, milk has the highest rate of discard intention at 68.3% followed by chicken at 41.3%, cereal at 40.9%, lettuce at 21.0%, and bread at 14.0%.

Figure 5. Mean Discard Rate by Product and Age



3.6.2. Logit Model for All Products

Table 12 shows the effect of label type and information on intended discard across all five products; fresh chicken, lettuce, bread, breakfast cereal, and milk. Since participants were given three packages of each product, three observations are recorded per participant

for each product for a total of nine observations per participant overall. For half of the participants, there are only six observations per participant due to one of the products, frozen chicken, being dropped from the analysis. The standard errors are clustered at the session level.

Three models are presented with differing combinations of the way product age is considered. It is important to note that product age is not the actual age of the product, but the age that is indicated by the date printed on the package by the research team. In the first model, product age is coded as the three ages used in the experiment for both the main effect by label type and interactions by label type. In the second model, the effect of product age with information is considered by if the package has a past-date label on it. Finally, in the third model, the effect of product age is only considered by if it is past date for all interactions. All three models control for individual characteristics such as participant age, income, gender, and children present in the household.[24] Participant characteristics are insignificant except for education. Those who hold four-year college degrees are more likely to discard products than those who do not hold four-year college degrees, but this effect is only significant at the 10% level. Other characteristics such as household size, kids, gender, income, and age of the participant are insignificant.

[24] Results with all participant characteristics included can be found in Appendix F.

Table 12. Effect of Label and Information on Consumer Discard Intention

Dependent variable: Discard Intention

	(1)		(2)		(3)	
	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect
Label Type						
Status Quo	-0.771**	-0.100	-0.771**	-0.100	-0.770**	-0.105
	(0.351)		(0.351)		(0.350)	
Standardized	0.114	0.015	0.114	0.015	0.112	0.015
	(0.331)		(0.331)		(0.332)	
Product Age - Medium						
SQ * Medium	1.420***	0.184	1.303***	0.169		
	(0.368)		(0.380)			
Standardized * Medium	0.450*	0.058	0.311	0.040		
	(0.230)		(0.229)			
Product Age - Old						
SQ * Old	2.369***[A]	0.307	2.441***	0.317		
	(0.395)		(0.385)			
Standardized * Old	1.441***[A]	0.187	1.519***	0.197		
	(0.230)		(0.207)			
Product Age - Post date						
SQ * Post date					1.945***	0.265
					(0.373)	
Standardized * Post date					1.005***	0.137
					(0.195)	
Information						
SQ * Info	-0.653[B]	-0.085	-0.653[D]	-0.085	-0.652[F]	-0.089
	(0.568)		(0.569)		(0.568)	
Standardized * Info	-0.498[B]	-0.065	-0.498[D]	-0.065	-0.496[F]	-0.068
	(0.355)		(0.355)		(0.354)	
Information * Product Age						
SQ * Info * Medium	0.113[C]	0.015				
	(0.657)					
Standardized * Info * Medium	0.497[C]	0.065				
	(0.395)					
SQ * Info * Old	0.525[C]	0.068				
	(0.631)					
Standardized * Info * Old	0.904**[C]	0.117				
	(0.353)					
SQ * Info * Post date			0.380[E]	0.049	0.394[G]	0.054
			(0.622)		(0.617)	
Standardized * Info * Post date			0.755**[E]	0.098	0.735**[G]	0.100
			(0.334)		(0.329)	
Constant	-2.511**		-2.505**		-2.485**	
	(1.193)		(1.194)		(1.142)	
Pseudo R^2	0.115		0.114		0.079	
Observations	2,679		2,679		2,679	

The same superscript letter indicates the estimates between label types are not significantly different at the 5% level as determined by a Wald test.
*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the session level.

In all three models, the excluded label type is *no label*. Participants received products with no labels — and thus no dates — printed on the package. Therefore, the estimates for the two label types, status quo and standardized, represent the effect of the label with a *young* or *pre-date* product age indicated on it in relation to a product with no label. Since only one of the product ages chosen for each product is pre-date, *young* and *pre-date* packages are identical. Status quo labels have a negative and significant effect on intended discard translating to a nine-percentage point decrease in discard when an unlabeled product is affixed with a young status quo label. In addition, a Wald test reveals that the effect of the status quo label is significantly different from the effect of the standardized label at the 5% level ($\chi^2(1) = 4.15$, $p = 0.042$). The presence of a label on a package decreases the probability of discard by about 10.0 and 1.5 percentage points for status quo labels and standardized labels, respectively; however, the standardized label's effect is not significantly different from zero. This does not hold true for other product ages with a standardized label though.

In Model 3, the effect of post-date labels on the intended discard is positive and significant across models thus supporting the findings of Roe et al. (2018). However, this effect is not purely from the product having a post-date label as the first two models indicate that there is a significant difference between the effects of medium age and old age for both status quo labels ($\chi^2(1) = 56.01$, $p = 0.000$) and standardized labels ($\chi^2(1) = 41.76$, $p = 0.000$).[25] Packages with medium-age status quo labels were discarded at a rate 16.9 percentage points higher than that of young-age (pre-date) packages with status quo labels.

[25] Results of Wald test from Model 2 estimates are reported.

Medium-age standardized labels increased intended discard by 4.0 percentage points over young-age packages with standardized labels. Old-age products were about 19 to 32 percentage points more likely to be discarded than young-age products of the same label type.

Across all three models, there is a significant difference in the effect of product age between status quo and standardized labels apart from the effect of old age in Model 1 ($\chi^2(1) = 3.81$, $p = 0.051$). Initially, these results suggest that standardization consistently has less of an effect on intended discard than status quo. However, for Models 1 and 2, the effects of medium and old age as reported in Table 12 are in relation to the main effect of that label type, e.g. the young-age labeled product of that label type. To compare the effect of product age across label type, the overall effect at each labeled product age needs to be calculated.

Table 13 presents the overall effect of each age by label type for the specifications from Models 1 and 2. As discussed above, there is a significant difference between status quo and standardized labels for young-age labels, but this difference disappears for medium- and old- age labels. This indicates that switching to standardized phrases without adjusting the date on the label could increase food waste since the rate of discard for status quo labels is significantly less than that of standardized labels for pre-date labels and no different for post-date labels; however, this consumer behavior could be changed through an information campaign.

Table 13. Effect of Product Age

Model:	(1)		(2)	
	Logit	Marginal Effect	Logit	Marginal Effect
Status Quo				
Young	-0.771**	-0.100	-0.771**	-0.100
	(0.351)		(0.351)	
Medium	0.650***^{I}	0.084	0.533**^{III}	0.069
	(0.248)		(0.224)	
Old	1.599***^{II}	0.207	1.670***^{IV}	0.217
	(0.232)		(0.242)	
Standardized				
Young	0.114^{A}	0.015	0.114^{B}	0.015
	(0.331)		(0.331)	
Medium	0.564*^{A,I}	0.073	0.425^{B,III}	0.055
	(0.327)		(0.307)	
Old	1.555***^{II}	0.202	1.633***^{IV}	0.212
	(0.276)		(0.275)	

The same superscript letter indicates the estimates between ages are not significantly different at the 5% level as determined by a Wald test.
The same superscript Roman numeral indicates the estimates between labels are not significantly different at the 5% level as determined by a Wald test.
*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the session level.

In Table 12, the effect of information is captured by the interaction terms between information and the label type since no participant who received unlabeled products also received label information. Therefore, for example, the estimate on status quo interacted with label information captures the difference between the discard intentions of the status quo group that received no information and the status quo group that was given information. It is important to remember that the information given to those who saw the status quo labels was different than the information given to those who saw the standardized labels. The status quo group received information about how the phrases on the labels indicated when the manufacturer believed "the taste and quality of their product [would] begin to decline" and that these dates were "chosen by the company and not regulated by any government rule or scientific agency." The standardized group received

information about the two specific phrases used in the two-phrase labeling system: “Best if used by” and “Use by.” The participants were told that these two phrases have specific meanings with “Best if used by” indicating quality and “Use by” indicating safety. The information given with the standardized labels also included the clause about the dates being unregulated. The main effect of information is negative, indicating that participants who received information were less likely to discard products, for both label types and across all three models; however, this effect, although consistent in magnitude, is not significant.

While the main effect indicates that information does not affect consumers’ discard intentions, there is the possibility that the information only affects consumer behavior for certain product ages. The final variable of interest in this model investigates this effect for post-date labels and each product age. Since these product age interactions are included, the insignificant estimates from the previous paragraph show that information does not influence participants’ discard decisions for young-age or pre-date products. Model 1 allows the effect of information on the interpretation of the date on the label to vary across the three product ages while Models 2 and 3 only allow the same effect to vary by whether the label is pre- or post-date. Despite these differences, the results consistently indicate that, all else equal, information does not have an effect on discard intention regardless of product age and label type. There is a significant effect on the old-age and post-date standardized interactions with information but when the overall effect for information at these ages is calculated, neither the effect at old-age ($\chi^2(1) = 2.37$, $p = 0.123$) or at post-date ($\chi^2(1) = 0.87$, $p = 0.351$) is significantly different from zero. In light of these findings, the interaction between information and product age is dropped for the remaining analysis.

3.6.3. Logit Model by Product

Table 14 presents the results of the logit regression by product with the accompanying marginal effects. The subsamples consist of the decisions made for each product. In each subsample, there are three observations per participant and only half the number of participants as the total sample since participants only saw one of two sets of products. The lettuce subsample is dropped due to insufficient variation of consumer behavior across the covariates for that product, i.e., only one participant intended to discard labeled, pre-date lettuce samples which led to convergence issues during estimation. The results from the remaining subsamples of products along with the full sample are presented. Since the results from the initial models indicated that information does not affect the discard decision differently by product age, the interaction term is dropped.

The first column displays this model with the full sample. As before, there is a significant difference between the product ages with pre- and post-date labels leading to a decrease and increase in the probability of discard, respectively. The difference between all three ages is significant indicating that the consumers take into account the actual date and not just whether the label is pre- or post-date. Status quo label information has a slight negative effect on intended discard, while the effect of information on standardized labels is not significant. The effect of information across the product subsamples is insignificant for both types of information.

Table 14. Effect of Label and Information on Consumer Discard Intention by Product

Dependent variable: Discard Intention	All Products		Chicken		Bread		Cereal		Milk	
	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect
Label Type										
Status Quo	-0.909***A	-0.118	-0.678C	-0.088	-0.611G	-0.034	0.615K	0.075	-1.378***	-0.245
	(0.296)	(0.039)	(0.513)	(0.066)	(0.969)	(0.055)	(1.075)	(0.131)	(0.508)	(0.090)
Standardized	-0.182A	-0.024	0.190C	0.025	-0.022G	-0.001	0.162K	0.020	-0.103	-0.018
	(0.367)	(0.048)	(0.813)	(0.104)	(1.026)	(0.057)	(0.985)	(0.120)	(0.504)	(0.090)
Product Age - Medium										
SQ * Medium	1.458***	0.190	2.433***D	0.315	-0.712H	-0.039	0.893L	0.109	1.552***	0.276
	(0.270)	(0.035)	(0.717)	(0.084)	(1.294)	(0.072)	(0.556)	(0.068)	(0.429)	(0.073)
Standardized * Medium	0.682***	0.089	1.546***D	0.200	0.000H	-0.000	1.600***L	0.195	-0.503	-0.089
	(0.230)	(0.029)	(0.447)	(0.060)	(0.512)	(0.028)	(0.530)	(0.067)	(0.336)	(0.059)
Product Age - Old										
SQ * Old	2.596***	0.338	3.136***E	0.406	1.504*I	0.083	2.469***M	0.301	3.158***	0.561
	(0.267)	(0.039)	(0.673)	(0.077)	(0.820)	(0.050)	(0.538)	(0.069)	(0.455)	(0.067)
Standardized * Old	1.886***	0.245	2.483***E	0.321	1.563***I	0.087	2.483***M	0.302	1.952***	0.347
	(0.217)	(0.028)	(0.551)	(0.073)	(0.493)	(0.027)	(0.539)	(0.065)	(0.291)	(0.044)
Information										
SQ * Info	-0.309*B	-0.040	-0.348F	-0.045	-0.155J	-0.009	-0.339	-0.041	-0.320N	-0.057
	(0.172)	(0.023)	(0.276)	(0.036)	(0.526)	(0.029)	(0.233)	(0.030)	(0.381)	(0.068)
Standardized * Info	0.121B	0.016	-0.030F	-0.004	-0.345J	-0.019	0.489	0.060	0.184N	0.033
	(0.264)	(0.034)	(0.655)	(0.085)	(0.664)	(0.038)	(0.377)	(0.044)	(0.372)	(0.066)
Demographics										
4-year Degree	0.420**	0.055	1.034***	0.134	1.165**	0.065	-0.037	-0.004	-0.039	-0.007
	(0.200)	(0.025)	(0.299)	(0.037)	(0.543)	(0.033)	(0.414)	(0.050)	(0.253)	(0.045)
Female	0.045	0.006	0.093	0.012	-0.373	-0.021	0.245	0.030	0.684**	0.121
	(0.168)	(0.022)	(0.263)	(0.034)	(0.549)	(0.030)	(0.384)	(0.046)	(0.348)	(0.064)
Constant	-2.496**		-6.880**		0.383		-2.368		-4.484**	
	(1.196)		(3.486)		(2.818)		(2.639)		(2.077)	
Pseudo R^2	0.112		0.193		0.114		0.175		0.178	
Observations	2679		537		537		534		534	

The same superscript letter indicates the estimates between label types for that category are not significantly different at the 5% level as determined by a Wald test.
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the session level.

Chicken is a unique product as it is the only product tested whose consumption is associated with actual and perceived safety concerns. When only one day remains on the label, the rate of discard is not significantly different than of an unlabeled chicken package; however, once the date on the label has passed, the rate of discard for labeled chicken increases significantly. For the status quo label, the increase in the rate of discard appears to slow after the date has passed with a 31.5 percentage point increase between one day pre-date and one day post day and only a 9.1 percentage point increase between one- and three-days post-date ($\chi^2(1) = 4.13$, $p = 0.042$). The strength of this relationship does hold for standardized labels, though, with corresponding 20.0 and 12.1 percentage point increases ($\chi^2(1) = 1.67$, $p = 0.196$). To further analyze the role of product age, the overall effect for each age is calculated.

Table 15 shows the overall effect of product age by label type on the discard rate as compared to the discard rate of unlabeled packages. While the effect of product age is significantly different between the three ages for both label types, the effect of product age is not significantly different between the status quo and standardized labels at all three ages tested indicating that consumers may only be paying attention to the date and not the associated phrase or they may be interpreting the two phrases in the same way. Either reason is concerning as the standardized phrase for chicken is "Use by," and manufacturers who adopt this language would likely affix a later date on the package than they would when using the phrase "Sell by." Since the effect of information for both label types is insignificant, a basic information campaign would be unlikely to affect consumer behavior. Knowing this, manufacturers could be conservative in adjusting the date associated with

the standardized label; however, this would defeat the purpose of a standardized labeling system.

Milk is the only product tested besides chicken that carries a "Sell by" phrase for the status quo label. As shown in Table 14, this phrase again leads to a decrease in intended discard, but, in the case of milk, this decrease is significant and much larger in magnitude. This could be partially or fully explained by the difference in days remaining on the label associated with each product; the young-age milk had two days remaining on the label while the young-age chicken only had one day remaining. The comparison between the standardized and status quo labels for milk differs from chicken as the standardized label phrase for milk is "Best if used by" as opposed to the "Use by" phrase on chicken. This difference is because, while consumers perceive safety concerns with milk, the consumption of post-date milk poses no actual safety concerns due to its pasteurization. Consumers were slower to react to older product age when the standardized phrase, "Best if used by," was on the label instead of the status quo phrase, "Sell by," so much so that there is no significant difference between the intended discard for young- (two days pre-date) and medium-age (one day post-date) standardized labeled milk ($\chi^2(1) = 2.23$, $p = 0.135$). Milk that is a week beyond its labeled date did show a significant increase in intended discard from the other product ages for both label types.

The overall effect of product age for milk is shown in Table 15. The effect of young- and medium-age labels vary between status quo and standardized labels with status quo labels having a 24.5 percentage point decrease in intended discard as compared to 1.8 for standardized labels. For medium age, both label types have effects that are insignificant from unlabeled, but significantly different from each other with the rate of discard for status

quo now higher than standardized. While the overall effect of the oldest product age is slightly lower for the status quo label at 31.6 percentage points versus 32.8 percentage points for the standardized label, the difference is insignificant ($\chi^2(1) = 0.06$, $p = 0.803$). This indicates that at one week past the date on the label, consumers are discarding the products labeled with "Best if used by" the same as products with "Sell by." This is again an interesting result as manufacturers would likely pair a later date with "Best if used by" than "Sell by," but once the day has passed, consumers don't appear to interpret the "Sell by" product as fresher than "Best if used by."

Bread and cereal both had the same status quo and standardized phrases: "Best before" for status quo and "Best if used by" for standardized. The effect of either label for bread did not appear until the label indicated it was three days past-date. Cereal showed a similar delay with the status quo label not having an effect until the label indicated the product was 70 days past-date while the standardized label increased discard once the label indicated the product was 10 days past-date. These effects are in relationship to the young-age product of the same label type though. The overall effect of product age in Table 15 for bread shows that only the old-age standardized label had a significant effect on intended discard in relation to unlabeled bread. Across each of the three product ages, there is no significant difference in effect between status quo and standardized labels for bread. This could be because the phrases are very similar and are paired with dates that only span four days or because consumers rely on the date label less with bread than other products. The insignificance of the phrase and product age points to the latter, although the narrow range of dates could also be driving this result.

While cereal has the same phrases for status quo and standardized labels as bread, the range of dates is much larger spanning 90 days. The overall effect of product age in Table 15 indicates that age does not have a significant effect on intended discard until the label is 70 days past-date. At this age, the status quo and standardized labels increase intended discard by 32.2 and 37.6 percentage points, respectively. As with all other products tested, this effect is not significantly different between the two label types ($\chi^2(1) = 1.26$, $p = 0.262$). In fact, cereal joins both bread and chicken in the list of products that show no significant difference between status quo and standardized labels for any of the three product ages tested.

Cereal is unique in that it is the only product of the four analyzed in which there was a significant difference in the effect of information between label types. The information given to those who received status quo labels decreased the rate of discard by 3.4 percentage points while the information on standardized labels increased discard by 8.9 percentage points for cereal. While neither of these were significantly different from those in their respective label groups who did not receive information, they were significantly different from each other at the 5% level ($\chi^2(1) = 4.23$, $p = 0.040$).

Table 15. Effect of Product Age by Product

	All Products		Chicken		Bread		Cereal		Milk	
	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect
Status Quo										
Young	-0.909***	-0.118	-0.678^{III}	-0.088	-0.611^{A,VI}	-0.034	0.615^{C,IV}	0.075	-1.378***	-0.245
	(0.296)		(0.513)		(0.969)		(1.075)		(0.508)	
Medium	0.549**^{I}	0.071	1.755***^{IV}	0.227	-1.324^{A,VII}	-0.073	1.508^{C,V}	0.184	0.174	0.031
	(0.216)		(0.577)		(1.072)		(1.053)		(0.332)	
Old	1.687***^{II}	0.220	2.458***^{V}	0.318	0.893^{VIII}	0.050	3.084***^{VI}	0.376	1.780***^{VII}	0.316
	(0.234)		(0.534)		(0.547)		(1.021)		(0.426)	
Standardized										
Young	-0.182	-0.024	0.190^{III}	0.025	-0.022^{B,VI}	-0.001	0.162^{IV}	0.020	-0.103^{D}	-0.018
	(0.367)		(0.813)		(1.026)		(0.985)		(0.504)	
Medium	0.500*^{I}	0.065	1.736***^{IV}	0.224	-0.022^{B,VII}	-0.001	1.762*^{V}	0.215	-0.606^{D}	-0.108
	(0.301)		(0.589)		(0.979)		(0.969)		(0.423)	
Old	1.704***^{II}	0.222	2.673***^{V}	0.346	1.541*^{VIII}	0.085	2.644***^{VI}	0.322	1.849***^{VII}	0.328
	(0.272)		(0.552)		(0.806)		(0.968)		(0.454)	

The same superscript letter indicates the estimates between ages are not significantly different at the 5% level as determined by a Wald test.
The same superscript Roman numeral indicates the estimates between between labels are not significantly different at the 5% level as determined by a Wald test.
*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the session level.

3.6.4. Food Management

In this final section of results, I focus on the socially optimal outcome of food management. In the previous models, only intended discard was examined. This does not account for the fact that it is socially optimal for some food to be discarded. Food that is unsafe to consume due to high health risks should not be consumed as the expected cost of the negative health effects would outweigh the expected benefits of consuming the food instead of discarding it. I introduce *mismanage* as the dependent variable in this final phase. Since all products in this study were "in-date" – products that had not reached the date originally printed on the label by the manufacturer – any discard would be considered waste except for one category, fresh chicken. There are two phrases in the standardized label system: "Best if used by" and "Use by." The "Use by" date is an indication of safety and is only used on products associated with increased health risks as the product ages. The only product in the study that falls into this category is fresh chicken. Since it is socially optimal for chicken with the standardized label to be discarded if it is past its printed date, *mismanage* is coded as 0 when discard is chosen for past-date standardized labeled chicken. If chicken with the standardized label that is past its printed date is not discarded by the consumer, then *mismanage* is coded as 1. For all other categories, *mismanage* is coded as 1 if the consumer intends to discard it and 0 if the consumer intends to keep it. Since this distinction only affects chicken out of the products tested in this study, the analysis is only performed with the fresh chicken subsample.

Table 16. Effect of Safety Phrase on Food Management (Chicken)

Dependent variable:	Discard		Mismanage	
	Logit	Marginal Effect	Logit	Marginal Effect
Label Type				
Status Quo	-0.678[A]	-0.088	-0.458[E]	-0.061
	(0.513)		(0.518)	
Standardized	0.190[A]	0.025	0.467[E]	0.062
	(0.813)		(0.626)	
Product Age - Medium				
SQ * Medium	2.433***[B]	0.315	2.403***[F]	0.319
	(0.717)		(0.709)	
Standardized * Medium	1.546***[B]	0.200	3.786***[F]	0.503
	(0.447)		(0.593)	
Product Age - Old				
SQ * Old	3.136***[C]	0.406	3.080***[G]	0.409
	(0.673)		(0.667)	
Standardized * Old	2.483***[C]	0.321	2.870***[G]	0.381
	(0.551)		(0.478)	
Information				
SQ * Info	-0.348[D]	-0.045	-0.267[H]	-0.035
	(0.276)		(0.304)	
Standardized * Info	-0.030[D]	-0.004	-0.428[H]	-0.057
	(0.655)		(0.469)	
Demographics				
4-year Degree	1.034***	0.134	0.232	0.031
	(0.299)		(0.301)	
Female	0.093	0.012	0.330	0.044
	(0.263)		(0.316)	
Constant	-6.880**		-4.829**	
	(3.486)		(1.911)	
Pseudo R^2	0.193		0.301	
Observations	537		537	

The same superscript letter indicates the estimates between label types are not signifincantly different at the 5% level as determined by a Wald test.
*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the session level.

The results of social optimum model for chicken are presented in Table 16. The discard model is included for comparison. In the social optimum model, the main effects of both label types are similar to those yielded by the discard model. The difference between the models is realized in the effect of product age once the product is past-date.

Recall that a positive coefficient means participants are more likely to mismanage their food by making an inappropriate decision (discard a pre-date or keep a past-date standardized product). This mismanagement is significantly higher for the standardized labels than the status quo labels when the date label is one day past ($\chi^2(1) = 22.51$, $p = 0.000$) but not when it is three days past ($\chi^2(1) = 3.25$, $p = 0.072$).

The results from the previous section provided evidence that the effect of product age for chicken on discard intention is linear; however, as shown in Table 17, the effect of product age with standardized labels on mismanagement appears to be nonlinear with medium-age standardized labeled chicken being the most mismanaged and significantly higher than old-age standardized chicken ($\chi^2(1) = 12.85$, $p = 0.000$) at 56.5 versus 44.3 percentage points more likely to be mismanaged than unlabeled packages. For chicken, medium age is one day past the date on the label while old age is three days past. The guidelines for standardized labeling state that any product with a safety phrase ("Use by") should be discarded after the date on the label. These results indicate that consumers, even those receiving information about the meaning of the standardized label phrases, are waiting beyond the date on the label to discard the product. Once the date is far enough in the past, consumers may be discarding the chicken purely because of the age, thus unintentionally avoiding mismanagement of the chicken. Consumers may lack the knowledge or understanding of standardized label interpretation or they may be in the habit of consuming chicken within a couple of days of the date on the label and not notice the change in the accompanying phrase. Since both estimates of label type interacted with information are insignificant, as has been the trend throughout all subsamples, a basic information campaign is unlikely to impact this consumer behavior.

Table 17. Effect of Product Age on Mismanagement (Chicken)

Dependent Variable	Discard		Mismanage	
	Logit	Marginal Effect	Logit	Marginal Effect
Status Quo				
Young	-0.678[I]	-0.088	-0.458[IV]	-0.061
	(0.513)		(0.518)	
Medium	1.755***[II]	0.227	1.946***	0.258
	(0.577)		(0.543)	
Old	2.458***[III]	0.318	2.622***[V]	0.348
	(0.534)		(0.518)	
Standardized				
Young	0.190[I]	0.025	0.467[IV]	0.062
	(0.813)		(0.626)	
Medium	1.736***[II]	0.224	4.253***	0.565
	(0.589)		(0.758)	
Old	2.673***[III]	0.346	3.337***[V]	0.443
	(0.552)		(0.725)	

The same superscript letter indicates the estimates between ages are not significantly different at the 5% level as determined by a Wald test.
The same superscript Roman numeral indicates the estimates between labels are not significantly different at the 5% level as determined by a Wald test.
*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the session level.

3.7. Conclusion

Food waste due to consumer confusion surrounding the safety and quality of products is prevalent. Without a label that gives some indication of the freshness of a product, I find that consumers intend to discard about 10% of high quality, safe, fresh food. Of the products tested, when consumers had to rely on their senses, milk was the most likely to be discarded with 31.7% being discarded, followed by lettuce at 7.4%, fresh chicken at 5.7%, and finally, bread and breakfast cereal at 4.6% and 3.7%, respectively. Across every label type and product age category, milk was consistently the most discarded product.

Providing a label on products gives consumers information on the freshness, quality, or safety of the product; however, there are no regulations or industry standards on how these labels should be set or the interpretation of a label. Consumers may not interpret the label in the way manufacturers intended for it to be interpreted or used. Retailer-facing labels such as those containing "sell by" are often used by consumers as such labels are found in clearly visible areas and are the only date label provided on the whole package within some product categories. In response to consumer confusion around date labels, a standardize two-phrase date labeling system has been proposed. This essay analyzes the effect of this standardization on intended discard of food and proper food management. It also considers an information campaign on date labeling in its current form and an information campaign on the proposed standardized two-phrase date labeling system.

When all five products studied are pooled, the overall effect of standardized labeling is similar to the status quo labeling except for when products are in-date. At this age, all else equal, products with a standardized date label are more likely to be discarded than those with a status quo label. This indicates that switching to standardized labeling could actually increase the amount of food waste if manufacturers switch only the phrase on the label. While the proposed standardized labeling system gives manufacturers definitions for the phrases, there is still no industry standard as to the appropriate time horizon to set the date label by for each product. This essay does not investigate how manufacturers set these dates in either the current labeling environment or the standardized labeling environment. However, it is often in the best interest of manufacturers to be conservative with these time horizons so that consumers do not experience an unappealing product before the date on the label. If manufacturers continue to conservatively date

packages while adopting the standardized phrasing, this could lead to a net increase in food being discarded by the consumer. In addition, conservative labeling would defeat the purpose of having a standardized date labeling system. If the dates do not accurately indicate the quality or safety decline they are supposed to, consumers will be in the same confused state in which they currently reside. Future work on manufacturers' procedures and implementation of a standardized date labeling system is needed.

Examining the effect of standardization by product provides insight into which paths labeling could be affecting consumer behavior. For three of the four products analyzed, chicken, bread, and cereal, the difference between status quo and standardized labels at each product age is insignificant. While this is not a surprising result for bread or cereal due to the similar nature of their status quo and standardized phrases, the similarity of effects between phrases for chicken is concerning. As discussed briefly in the results, manufacturers would need to affix a later date to a label with "Use by," the standardized phrase for chicken, than a "Sell by" label, the status quo phrase for chicken, in order to comply with the standardized definition of "Use by." However, these results indicate that consumers do not behave differently based on the phrase. This could be due to consumers only referring to the date on the label and ignoring the phrase, consumers ignoring any phrase in the current unstandardized environment, or consumers lacking knowledge about what the standardized phrase "Use by" means in the standardized environment.

The first two reasons could be the result of the unstandardized date labeling environment. If consumers are aware that the phrase does not mean anything, then they may use the date as part of a "rule" they have learned from others or through experience with the product in the current unstandardized labeling environment. They could

knowingly or unknowingly be ignoring the phrase. This problem could be magnified by the fact that the phrase printed on labels are usually smaller than the date itself. This increases the likelihood of consumers unknowingly ignoring the phrase through raising the search cost. The current unstandardized environment also lowers the benefit of searching for this phrase since phrase interpretation has no current industry standard.

The third reason for the effect of product age being unaffected by label type is that consumers lack knowledge about what the standardized phrase means. The effect of standardized information was insignificant, but the information was brief and given shortly before participants were asked to make discard decisions. A more intense information campaign along with a consistent standardize labeling environment where consumers gain experience with standardized labels could alter the way the standardized phrase is interpreted compared to the status quo phrase. Information campaigns could appear as part of larger food waste reduction education and serve as a marketing tool for manufacturers.

As an often-misunderstood product and the most discarded product in the experiment, milk yielded unique results that warrant their own discussion. Across product ages, the effect of the status quo label, "Sell by," and the standardized label, "Use by," on intended discard varied. For young-age cartons, those two days pre-date, the effect of product age was significantly lower for status quo than standardized label phrases. When the age moves to one day post-date, the effect flipped so that the effect of product age was significantly higher for status quo than standardized label phrases. This is unexpected as "Sell by" compared to "Best if used by" would generally be used to indicate a fresher product. However, this effect could be due to consumers reacting to the ambiguity of "Sell by." Since milk is a product often associated with safety risks – even though pasteurization

mitigates safety risks – consumers faced with an ambiguous label might react with caution by discarding the product. Once the milk is seven days past-date, consumers discard it at the same rate as milk with the standardized label, "Best if used by." At this age, a possible ambiguity effect could be taken over by the effect of the date being so far past.

A limitation of this study is the number of product ages that are used. Only three ages are tested for each product, and all are clustered around the label date. Variation in pre-date and post-date labels could provide more insight into the role of product age versus the product being in-date or out-of-date. The results from this experiment should not be extrapolated to dates beyond the range studied, and even interpolation should be done with caution given the evidence of nonlinearity among the age effects.

The final section of this study examines the effectiveness of date label standardization in achieving socially optimal consumer behavior regarding foods with safety risks associated with their consumption. Chicken was the only product in the experiment that fell into this category. The results from the discard model yield insignificant differences in the effect of product age between the status quo label "Sell by" and the standardized label "Use by." When mismanagement is considered instead of discard behavior, the results shift revealing an increase in suboptimal consumer behavior when the chicken is one day past-date. This increase indicates that, even with the "Use by" phrase, consumers may keep potentially unsafe foods beyond the date on the label. The effect is lessened at three days past-date and is not significantly different from the status quo label, but this decrease could be because of the date rather than the phrase.

Increased consumption of unsafe foods could be a serious unintended consequence of switching to a standardized date labeling system if not addressed as part of the

implementation of the standardized system. This is another reason manufacturers, concerned for their reputation and their customers' health, might resort to conservative time horizons. A key component of the standardized date labeling system is for manufacturers to base the time horizon for their products on either quality or safety (depending on the product), and not consumer behavior. If this manufacturing decision is not made independently from consumer behavior, confusion surrounding date labels will persist.

Bibliography

Ahern, J. J. (2006). The unintended consequences of a ban on the humane slaughter (processing) of horses in the United States (Doctoral dissertation, University of California, Davis, California).

American Association of Equine Practitioners. (2005). Hosted Unwanted Horse Summit. The American Horse Council Annual Meeting; Washington, DC.

Aschemann-Witzel, J. (2018). Consumer perception and preference for suboptimal food under the emerging practice of expiration date based pricing in supermarkets. *Food Quality and Preference*, 63, 119-128.

Bechert, U., Bartell, J., Kutzler, M., Menino, A., Bildfell, R., Anderson, M., & Fraker, M. (2013). Effects of two porcine zona pellucida immunocontraceptive vaccines on ovarian activity in horses. *The Journal of Wildlife Management*,77(7), 1386-1400.

BLM. (2019, March 1). *Program Data*. Retrieved from https://www.blm.gov/programs/wild-horse-and-burro/about-the-program/program-data

Bloom, J. (2010). *American wasteland: How America throws away nearly half of its food (and what we can do about it)*. Da Capo Lifelong Books.

Borra, S., Fikes, D., & Thesmar, H. (2017). U.S. Grocery Shopper Trends 2017: Food retail implications for U.S. Grocery Shopper Trends, Webinar, July 18, 2017, https://www.fmi.org/docs/default-source/webinars/trends-2017-webinar-7-18-2017.pdf: FMI.

Carson, R. T., & Czajkowski, M. (2018). A new baseline model for estimating willingness to pay from discrete choice models(No. 2018-04).

Council, N. R. (2013). Using Science to Improve the BLM Wild Horse and Burro Program. *A Way Forward*.

De Steiguer, J. E. (2011). *Wild horses of the West: history and politics of America's mustangs*. University of Arizona Press.

Dines, L. (2001). *The American Mustang Guidebook: History, Behavior, and State-By-State Directions on Where to Best View America's Wild Horses*. Willow Creek Press.

Elizondo, V., Fitzgerald T., & Rucker, R. (2016). You Can't Drag Them Away: An Economic Analysis of the Wild Horse and Burro Program. *Journal of Agricultural and Resource Economics*,41(1):1-24.

Food and Agriculture Organization of the United Nations. (2013*).* Food wastage footprint: impacts on natural resources: summary report. FAO.

GAO (U.S. Government Accountability Office). (2008). Effective Long-Term Options Needed to Manage Unadoptable Wild Horses. Washington, DC: U.S. Government Accountability Office.

GAO (Government Accountability Office). (2011). Horse Welfare: Action Needed to Address Unintended Consequences from Cessation of Domestic Slaughter. Report to Congressional Committees GAO-11-228. Washington, DC: U.S. Government Accountability Office.

Garrott, R. A., & Oli, M. K. (2013). A critical crossroad for BLM's wild horse program. *Science*, 341(6148), 847-848.

GMA & FMI. (2017). FMI – GMA Product Code Dating Initiative. Retrieved from https://www.gmaonline.org/file-manager/FINAL%20Fact%20Sheet%20-%20Product%20Code%20Dating%20Initiative%20-%20021417.pdf.

Godfrey, E. B., & Lawson, P. (1986). Wild horse management: An economic perspective. *Journal of Equine Veterinary Science*, 6(5), 266-272.

Gunders, D. (2012). Wasted: How America Is Losing Up to 40 Percent of Its Food from Farm to Fork to Landfill. NRDC.

Hall-Phillips, A., & Shah, P. (2017). Unclarity confusion and expiration date labels in the United States: A consumer perspective. *Journal of Retailing and Consumer Services*, 35, 118-126.

Hole, A.R. (2007). Fitting missed logit models by using maximum simulated likelihood. *The Stata Journal, 7*(3), 388-401.

Hole, A. R. (2008). Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment. *Journal of Health Economics*, 27(4), 1078-1094.

Kirkpatrick, J. F., Rutberg, A. T., Coates-Markle, L., & Fazio, P. M. (2012). Immunocontraceptive Reproductive Control Utilizing Porcine Zona Pellucida (PZP) in Federal Wild Horse Populations. Science and Conservation Center, Billings, Montana.

Koncel, M. A., & Rutberg, A. T. (2012). Catching the spirit: A study of bureau of land management wild horse adopters in New England. *Journal of Applied Animal Welfare Science*, 15(1), 32-52.

Lansford, N. H., Freeman, D. W., Topliff, D. R., & Walker, O. L. (1998). Hedonic pricing of race-bred yearling Quarter Horses produced by Quarter Horse sires and dams. *Journal of Agribusiness*, 16, 169-186.

Leib, E. B., Rice, C., Neff, R., Spiker, M., Schklair, A., & Greenberg, S. (2016). Consumer perceptions of date labels: National survey. Safety, 23(54), 19.

Lenz, T. R. (2009). The unwanted horse in the United States: An overview of the issue. *Journal of Equine Veterinary Science*, 29, 253–258.

Lipinski, B., Clowes, A., Goodwin, L., & Hanson, C. (2017). *SDG Target 12.3 on Food Loss and Waste: 2017 Progress Report*. Washington, DC: Champions, 12.

Mask, T. A., Schoenecker, K. A., Kane, A. J., Ransom, J. I., & Bruemmer, J. E. (2015). Serum Antibody immunoreactivity to equine zona protein after SpayVac vaccination. *Theriogenology*, 84(2), 261-267.

Maynard, L. J., & Stoeppel, K. M. (2007). Hedonic price analysis of thoroughbred broodmares in foal. *Journal of Agribusiness*, 25(2), 181.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447-470.

Mercier, S., Villeneuve, S., Mondor, M., & Uysal, I. (2017). Time–temperature management along the food cold chain: A review of recent developments. *Comprehensive Reviews in Food Science and Food Safety*, 16(4), 647-667.

Neibergs, J. S. (2001). A hedonic price analysis of Thoroughbred broodmare characteristics. *Journal of Agribusiness*, 17(2), 299-314.

Newsome, R., Balestrini, C. G., Baum, M. D., Corby, J., Fisher, W., Goodburn, K. & Yiannas, F. (2014). Applications and perceptions of date labeling of food. *Comprehensive Reviews in Food Science and Food Safety*, 13(4), 745-769.

North, M. S., Bailey, D., & Ward, R. A. (2005). The potential impact of a proposed ban on the sale of US horses for slaughter and human consumption. *Journal of Agribusiness*, 23(1), 1-17.

Qi, D., & Roe, B. E. (2016). Household food waste: multivariate regression and principal components analyses of awareness and attitudes among US consumers. *PloS One*, 11(7), e0159250.

Quested, T. (2013). The Milk Model: Simulating Food Waste in the Home. WRAP.

Quested, T., & Murphy, L. (2014). Household food and drink waste: a product focus. United Kingdom: Banbury (UK): Waste and Resources Action Programme.

Ransom, J. I., Roelle, J. E., Cade, B. S., Coates-Markle, L., & Kane, A. J. (2011). Foaling rates in feral horses treated with the immunocontraceptive porcine zona pellucida. *Wildlife Society Bulletin*, 35(4), 343-352.

ReFED. (2016). *A roadmap to reduce U.S. food waste by 20 percent: Executive Summary*. ReFED.

Roe, B.E., Phinney, D.M., Simons, C.T., Badiger, A.B., Bender, K.E., Heldman, D.R. (2018). Discard intentions are lower for milk presented in containers without date labels. *Food Quality and Preference*, 66(December 2017), pp.13–18.

Sen, S., & Block, L. G. (2008). "Why my mother never threw anything out": the effect of product freshness on consumption. *Journal of Consumer Research*, 36(1), 47-55.

Stowe, C. J., & Kibler, M. L. (2016). Characteristics of Adopted Thoroughbred Racehorses in Second Careers. *Journal of Applied Animal Welfare Science*, 19(1), 81-89.

Swegen, A., & Aitken, R. J. (2014). Prospects for immunocontraception in feral horse population control: exploring novel targets for an equine fertility vaccine. *Reproduction, Fertility and Development*.

Taoukis, P.S., & Labuza, T.P. (2003). Time-temperature indications (TTIs). *Novel Food Packaging Techniques*, 103-126.

Taylor, M., & Sieverkropp, E. (2013). The Impacts of US Horse Slaughter Plant Closures on a Western Regional Horse Market. *Journal of Agricultural and Resource Economics*, 38(1), 48-63.

Thompson, B., Toma, L., Barnes, A. P., & Revoredo-Giha, C. (2018). The effect of date labels on willingness to consume dairy products: Implications for food waste reduction. *Waste Management*, 78, 124-134.

Train, K., & Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. *Applications of Simulation Methods in Environmental and Resource Economics* (pp. 1-16). Springer, Dordrecht.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Tsiros, M., & Heilman, C. M. (2005). The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories. *Journal of Marketing, 69*(2), 114-129.

United States Census Bureau. (2017, July 1). Quick Facts. Retrieved November 28, 2018 from https://www.census.gov/quickfacts/fact/table/US/PST045217.

U.S. Department of Agriculture. (December 16, 2016). USDA Revises Guidance on Date Labeling to Reduce Food Waste. Retrieved June 15, 2019 from https://www.fsis.usda.gov/wps/portal/fsis/newsroom/news-releases-statements-and-transcripts/news-release-archives-by-year/archive/2016/nr-121416-01.

U.S. Department of Interior, Bureau of Land Management. (1998). Adoption of Wild Horse and Burros. *BLM Handbook H-4750-2*, U.S. Department of Interior, Bureau of Land Management. Washington, DC.

U.S. Department of Interior, Bureau of Land Management. (2016). Herd Management and Herd Area Statistics. Available online at http://www.blm.gov/wo/st/en/prog/whbprogram/herd_management/Data.html

Verbeke, W., & Ward, R.W. (2006). Consumer interest in information cues denoting quality, traceability, and origin: An application of ordered probit models to beef labels. *Food Quality and Preference*, *17*(6), 453-467

Vickner, S. S., & Koch, S. I. (2001). Hedonic pricing, information, and the market for thoroughbred yearlings. *Journal of Agribusiness, 19*(2), 173-190.

Wilson, N. L., Rickard, B. J., Saputo, R., & Ho, S. T. (2017). Food waste: The role of date labels, package size, and product category. *Food Quality and Preference*, 55, 35-44.

Wilson, N. L., Miao, R., & Weis, C. (2018). Seeing Is Not Believing: Perceptions of Date Labels over Food and Attributes. *Journal of Food Products Marketing*, 24(5), 611-631.

Yiannas, Frank. (2019). Letter to the Food Industry: Date Labeling. FDA. Retrieved from https://www.fda.gov/media/125114/download

Yu, Y., & Jaenicke, E. C. (2019). The Effect of Sell-By Dates on Purchase Volume and Food Waste: A Case of New York City's Sell-By Regulation of Milk (No. 273862). Agricultural and Applied Economics Association.

Yu, Yang and Jaenicke, Edward C. (2018). The Effect of Sell-by Dates on Purchase Volume and Food Waste. Available at SSRN.

Appendix A. Milk Label Information

Milk Label Information

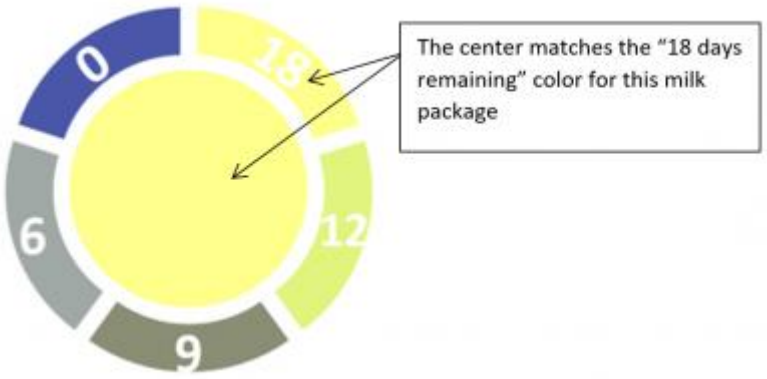
The rest of the containers of milk you will see and smell will have one of the following labeling formats:

Date Label

- This date is printed on the milk container by the company that puts the milk in the container
- The date indicates the company's best estimate of the day when the milk in the container will no longer be of a quality that most people will fully enjoy
- Currently the best estimate is that this date occurs 18 days after milk is put into the container
- This means that the company estimates that milk will have an 18-day shelf life after it is put into a container

Real-time Shelf-life Indicator (RSI) Label

- The RSI label indicates the freshness of milk in a new way.
- The center dot of the indicator changes color to show how many days the milk has remaining in its shelf life.
- The numbers surrounding the dot indicate the days remaining before end of shelf life (the color of these surrounding numbers do not change).



- If the milk is accidentally left out of the refrigerator, the label changes color more quickly to reflect that the milk's taste will likely turn bad more quickly
- This picture below shows the indicator changing over time as milk is stored in the refrigerator.

- Use this image as a guide for the rest of today's session.
- Compared to milk in containers with standard date labels, the milk in containers with this type of label:
    - come from the same farms and cows
    - go through the same milk processing plant
    - are treated and stocked the same way at the store

Appendix B. Milk Discard Results

Table 18. Logit Discard Results with Participant Characteristics

Dependent variable: Pr(Discard)	Model 1		Model 2	
	Logit	Marginal Effect	Logit	Marginal Effect
Milk Attributes				
Label	-0.134	-0.033	-0.139	-0.034
	(0.196)	(0.049)	(0.197)	
Label*Days Remaining	-0.053***	-0.013	-0.054***	-0.013
	(0.016)	(0.004)	(0.017)	
Quality Rank	1.057***	0.257	1.059***	0.257
	(0.291)	(0.063)	(0.293)	
Participant Characteristics				
Age	0.107***	0.026	0.099***	0.024
	(0.027)		(0.030)	
Age²	-0.001***	0.000	-0.001***	0.000
	(0.000)		(0.000)	
ln(income)	-0.229***	-0.056	-0.232***	-0.056
	(0.061)		(0.070)	
Rural	0.213	0.052	0.091	0.022
	(0.272)		(0.322)	
Suburban	0.219*	0.053	0.155	0.038
	(0.132)		(0.150)	
Four-year degree	0.119	0.029	0.106	0.026
	(0.195)		(0.193)	
Female	0.056	0.014	0.096	0.023
	(0.096)		(0.091)	
Lives alone			0.012	0.003
			(0.146)	
Household size			0.022	0.005
			(0.038)	
Purchases whole milk			-0.019	-0.005
			(0.150)	
Purchases milk by gallon			0.210	0.051
			(0.134)	
Purchases Kroger milk			0.050	0.012
			(0.121)	
Constant	-1.520		-1.482	
	(1.043)		(1.002)	
Psuedo R^2	0.068		0.069	
Observations	880		880	

*, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.
Standard errors in parentheses and clustered at the milk sample level.
Marginal effect at the mean reported.

Appendix C. WTP Results Across All Participants

Table 19. Mixed Logit Results Across All Participants

Variable	Estimates	SD	Willingness to Pay
Prob(Purchase=1)			
Price ($/gallon)	-3.086*** (0.582)		
Label Days Remaining	0.664*** (0.095)	0.286*** (0.065)	$0.22 [0.16, 0.27]
Days Remaining x Low Discard Sensitivity	-0.216* (0.121)		-$0.07 [-0.15, 0.01]
Days Remaining x High Discard Sensitivity	0.123 (0.156)		$0.04 [-0.06, 0.14]
RSI Label	-1.674*** (0.569)	-0.057 (0.405)	-$0.54 [-0.96, -0.13]
RSI x Young Age Category	1.077** (0.512)		$0.35 [0.00, 0.69]
RSI x Middle Age Category	0.810* (0.467)		$0.26 [-0.05, 0.57]
Status Quo	-7.754*** (1.612)	4.434*** (1.108)	-$2.51 [-3.15, -1.88]
Log Likelihood	-244.467		
AIC	512.934		
BIC	577.071		
Observations	516		516

*, **, *** indicates significance at the 10%, 5%, and 1% levels, respectively.
Marginal effects are reported in the first column along with standard errors in parentheses.
95% confidence intervals reported in square brackets.

Appendix D. Date Labeling Information

Status Quo Labeling Information

Today you will see several packages of food purchased from local stores. These packages have a date printed on them along with a phrase such as 'Best By', 'Sell By' or 'Best if Used By'.

- These dates are about food freshness and quality
- The company making this product thinks the taste and freshness of their product will decline after this date
- The date does not indicate product safety – products often remain safe to eat after this date so long as they look and smell appropriate
- The date is chosen by the company making the product and is not regulated by any government rule or scientific agency
- Fresh meat taken from a previously opened package may be safely eaten for 3-5 days after the package has been opened if the package was properly refrigerated and the product smells and looks appropriate

Standardized Labeling Information

Today you will see several packages of food purchased from local stores with a date printed on the label.

- If the label says **Best if Used By**:
    - The date is about product freshness
    - The company making this product thinks the taste and freshness of their product may decline after this date
    - 'Best if Used By' dates do not indicate product safety – products often remain safe to eat after this date so long as they look and smell appropriate
    - The date is chosen by the company making the product and is not regulated by any government rule or scientific agency
- If the label says **Use By**:
    - The date is about product safety
    - The company making this product thinks the product may be unsafe if consumed after the 'Use By' date
    - The date is chosen by the company making the product and is not regulated by any government rule or scientific agency
- Fresh meat taken from a previously opened package may be safely eaten for 3-5 days after the package has been opened if the package was properly refrigerated and the product smells and looks appropriate

Appendix E. Product Packaging

Fresh Chicken

Lettuce

Bread

Milk

Cereal


BEST IF USED BY
26 AUG 2019
BEST IF USED BY
14 NOV 2019


OAT CLUSTERS
OAT CLUSTERS
OAT CLUSTERS
OAT CLUSTERS
OAT CLUSTERS
OAT CLUSTERS

Appendix F. Logit Results

Table 20. Logit Results with Participant Characteristics

Dependent variable: Discard Intention

	(1)		(2)		(3)	
	Logit	Marginal Effect	Logit	Marginal Effect	Logit	Marginal Effect
Label Type						
Status Quo	-0.752**	-0.097	-0.752**	-0.098	-0.752**	-0.102
	(0.351)		(0.351)		(0.351)	
Standardized	0.130	0.017	0.130	0.017	0.128	0.017
	(0.335)		(0.335)		(0.335)	
Product Age - Medium						
SQ * Medium	1.421***	0.184	1.304***	0.169		
	(0.368)		(0.380)			
Standardized * Medium	0.450*	0.058	0.312	0.040		
	(0.230)		(0.229)			
Product Age - Old						
SQ * Old	2.371***	0.307	2.443***	0.317		
	(0.395)		(0.385)			
Standardized * Old	1.443***	0.187	1.522***	0.197		
	(0.230)		(0.207)			
Product Age - Post date						
SQ * Post date					1.946***	0.265
					(0.372)	
Standardized * Post date					1.007***	0.137
					(0.196)	
Information						
SQ * Info	-0.659	-0.085	-0.659	-0.086	-0.658	-0.090
	(0.567)		(0.567)		(0.567)	
Standardized * Info	-0.491	-0.064	-0.491	-0.064	-0.489	-0.067
	(0.356)		(0.356)		(0.356)	
Information * Product Age						
SQ * Info * Medium	0.112	0.015				
	(0.657)					
Standardized * Info * Medium	0.498	0.064				
	(0.395)					
SQ * Info * Old	0.524	0.068				
	(0.631)					
Standardized * Info * Old	0.904**	0.117				
	(0.353)					
SQ * Info * Post date			0.379	0.049	0.393	0.053
			(0.622)		(0.617)	
Standardized * Info * Post date			0.755**	0.098	0.735**	0.100
			(0.334)		(0.329)	
Participant Characteristics						
Young-age participant	-0.123	-0.016	-0.123	-0.016	-0.118	-0.016
	(0.236)		(0.236)		(0.225)	
Middle-age participant	-0.114	-0.015	-0.114	-0.015	-0.109	-0.015
	(0.192)		(0.191)		(0.183)	
Household size	-0.063	-0.008	-0.063	-0.008	-0.060	-0.008
	(0.050)		(0.050)		(0.048)	
Four-year degree	0.381*	0.049	0.381*	0.049	0.364*	0.050
	(0.199)		(0.199)		(0.191)	
ln(income)	0.024	0.003	0.023	0.003	0.022	0.003
	(0.115)		(0.115)		(0.110)	
Female	0.027	0.003	0.027	0.004	0.026	0.004
	(0.170)		(0.170)		(0.162)	
Female * Kids	0.396	0.051	0.396	0.051	0.379	0.052
	(0.357)		(0.358)		(0.342)	
Kids	-0.247	-0.032	-0.246	-0.032	-0.236	-0.032
	(0.335)		(0.336)		(0.322)	(0.043)
Constant	-2.523**		-2.517**		-2.496**	
	(1.216)		(1.219)		(1.167)	
Pseudo R^2	0.116		0.115		0.080	
Observations	2,679		2,679		2,679	