

**Cluster Analysis of Consumer Attitudes Towards Online Grocery Shopping and Impacts
on Online Grocery Usage and Food Expenditure**

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Harrison T. Clark

May 2024

Acknowledgments

I'd like to sincerely thank Dr. Xuqi Chen for the unwavering guidance and support he has provided throughout my master's program and research. Being under his mentorship has been an incredible honor, enriching and expanding my academic experience significantly. His expert advice and patience were crucial to my research and the completion of my thesis and numerous passion projects. The memories and experiences I've gained at the University of Tennessee are invaluable and will be forever cherished, and Dr. Chen will forever be the man who made this Master's degree possible for me as I feel my experience would not have been possible without his understanding, patience, and support of my chase of carving my unique path through the last two years.

My committee members, Dr. Jacqueline Yenerall and Dr. Andrew Muhammad have been pillars of support for me, offering invaluable guidance that illuminated my path toward completing this study. Your collective wisdom not only steered me through academic challenges but also instilled in me a deeper appreciation for our field's complexity and beauty. Your encouragement was a beacon that lit my way, for which I am eternally grateful. I know I have the support of these three humans for all of my time and they have not stopped making that evident since day one of the program.

I also want to thank the faculty and staff at the Department of Agricultural and Resource Economics for their educational support and assistance during my time here. A special mention goes out to Dr. Carlos Trejo-Pech, Dr. Edward Yu, Dr. Sreedhar Uppendrum, Dr. Alicia Rihn, Dr. William Johnson, Mr. Ty Wolaver, Mrs. Amy Cooley, Mrs. Brittany Gentry, Mrs. Julie Dougherty, Mr. Chris Rhodes, Mrs. Renee Sharp, and Ms. Ann Allen for their kindness and guidance throughout the program and for always pushing me forward and having my back. My

fellow graduate students were my rock through this with every step; I have a love for every one of them from the bottom of my heart and will always cherish the memories and relationships I take forward from this program, most notably my relationship with my brother Kumar Bolakhe. Never have I become so close with a group of peers in my life; the turbulence and challenge of this program ingrained us as one and have made us inseparable by distance or time, and this is one of the most valuable things I will walk away with.

Abstract

Online grocery shopping has exploded in prevalence since the COVID-19 pandemic, and literature has come out about best practices for marketing generally to the population to incentivize additional usage and investigate the attitudes that consumers hold towards online grocery shopping. In this study, we focus on not just general marketing but targeted online marketing strategies that will support the goals of firms within the online grocery store sector, by completing a detailed clustering analysis, followed by regression analysis to determine how the segmentation of consumers affected their willingness to use online grocery stores in the past six months and their expenditures both online and in-person. By analyzing these, we deepen our understanding of the current barrier of online grocery shopping, provide insights into consumer behavior online, and provide reference information to retailers in their mission to better set up consumer target group and optimize their goals.

Keywords: Consumer Behavior, Online Grocery, Clustering Analysis

Table of Contents

Chapter 1 Introduction and General Information	1
<i>Problem Identification</i>	<i>1</i>
<i>Objective Statement.....</i>	<i>4</i>
<i>Hypotheses.....</i>	<i>4</i>
Chapter 2 Literature Review	6
<i>Online Grocery Shopping Services.....</i>	<i>6</i>
<i>Home Delivery Services</i>	<i>6</i>
<i>Convenience and Time-Saving.....</i>	<i>7</i>
<i>Perceived Risk</i>	<i>8</i>
<i>Demographic Factors</i>	<i>9</i>
Chapter 3 Materials and Methods.....	11
<i>Data Collection.....</i>	<i>11</i>
<i>Survey Design.....</i>	<i>11</i>
<i>Data Analysis Method.....</i>	<i>12</i>
<i>Model Specification.....</i>	<i>14</i>
Chapter 4 Results and Discussion	16
<i>Cluster Analysis.....</i>	<i>16</i>
<i>Regression Analysis</i>	<i>27</i>
<i>Robustness Check</i>	<i>31</i>
Chapter 5 Conclusions and Recommendations.....	36
<i>Implications.....</i>	<i>36</i>
<i>Limitations and Future Research.....</i>	<i>42</i>
<i>Conclusion.....</i>	<i>43</i>
List of References	46
Vita	51

List of Tables

Table 4.1. Table of the Four Cluster Analysis.	Error! Bookmark not defined.
Table 4.2. Table of Summary Statistics of Relevant Variables for Regression..	Error! Bookmark not defined.
Table 4.3. Results of Logit and OLS Regressions.	.29
Table 4.4. VIF test for the regressions to test multicollinearity.....	33
Table 4.5. Regression for Robustness Check.....	34

Chapter 1

Introduction and General Information

Problem Identification

The online grocery store market in the United States has grown steadily over the past ten years (IBIS, 2023), as seen in the consistent development of new online grocery stores, with a 19.9% increase in the number of online grocery businesses from 2018 to 2023 (IBIS, 2023). From tech giants like Amazon Fresh, which operates in all 50 states, to regional grocers like Fresh Direct, which only operates in the Northeast, firms are cementing their roots across the nation. In 2017, online grocery retailers brought in 399.66 million dollars, which has skyrocketed to 925.36 million dollars in 2023 (Statista, 2024). With the industry's sales growth having no end in sight, the number of businesses developed to service the growing demand for online groceries will subsequently continue to increase. With an increase in market options for consumers, retailers must invest in practices that are empirically led and are shown to garner consumer attention by satisfying their preferences to remain competitive in gaining new users and increasing revenue from existing users.

Newcomers to online grocery shopping are likely to shop online with the retailer they are most familiar with, which leads to why Amazon and Walmart lead the pack as number one and number two top-used services in terms of net sales volume. Yet, as consumers become more familiar with online grocery shopping, they are likely to search outside of the familiar to find online options that better suit their preferences when purchasing groceries (Melis, 2015). While online retailers have made substantial efforts to entice newcomers to the market, only 16% of US adults have ordered groceries online (Saphores and Xu, 2020). Current consumers are drawn by the convenience of ordering a week's worth of groceries without taking a step in a store,

providing perceived time savings and reducing one's impulsivity of purchasing while shopping, but many fear that there are higher costs associated with shopping online for groceries, there isn't truly time saved through shopping online, that there are security issues with online retailers, or that without being able to touch and feel their produce they will be unsatisfied with what they receive. With only 31% of US consumers presently possessing any online grocery shopping experience, however, the potential for the growth of online grocery shopping is vast with a greater understanding of these barriers and how they affect different clusters of consumers (Etumnu, 2020).

While barriers to online grocery shopping have been extensively reviewed over the past twenty years, much of the findings are foundational understandings of the behavioral traits that online grocery consumers possess and their general sentiment towards online grocery store interfaces and traditional delivery systems. For example, (Huyghe, 2017) found that consumers perceive different types of online product displays in alternate manners, and (Harris-Lagoudakis, 2022) found that online grocery retailers' interface affects the purchasing of healthier food items. Moreover, consumers have been found to view different forms of online shopping (delivery versus pickup) offered by firms with different willingness to consume (van Ewijk, 2020).

Although the research on the consumer perception of online grocery store attributes has been conducted with great precision, the examination of consumer perception towards actual actions to address these perceptions is lacking. A better understanding of consumer attitudes and preferences, as well as the factors influential in consumers' decisions regarding the adoption and usage of online grocery shopping will enable retailers to better serve the needs of their customers and attract new customers.

While customers are attracted to the convenience of online grocery shopping, purchasing certain foods (e.g., produce) online is perceived as risky by many online grocery consumers (Mortimer, 2016). The lack of ability to control and know the sensory attributes of these products leaves consumers unwilling to purchase these products (Campo, 2015). An innovative solution to this issue is advanced packaging technology to prevent the quality degradation of produce through the delivery process. Although online grocery shopping does drive healthier basket-level consumption by reducing the number of impulse purchases (Huyghe, 2017), more progress could be made in advancing the welfare of online grocery consumers by developing methods to increase the consumption of healthy produce.

Assuring that consumers would be willing to purchase products under these new risk conditions and if they would be willing to pay a minimal fee for advanced packaging would offer two potential methods to increase the consumption of healthy produce. A minimal fee will remind the consumer that their products will arrive safely. Giving the option to use advanced packaging may provide consumers with an assured sense of control over the quality of their product purchases. All of these benefits to human health are simultaneously benefits to profits of those in the produce supply chain, as increased consumption will result in increased profits and resolution of much of the difficulty with selling produce online by grocery stores. Not only are these welfare gains feasible, but additionally this study will identify the impact of online grocery retailers beginning to accept SNAP benefits in recent years on those recipient's usage and expenditure within the OGS space. In the past, SNAP recipients have disinterest with shopping online for groceries for quality and financial reasons (Rogus et. al., 2019) but have shown interest in faster, more affordable online shopping experiences which has been cultivated through maturing of the industry for many retailers (Martinez et. al., 2018). With the fiscal collaboration

with retailers in recent years to expand SNAP benefits acceptance at the quickly maturing and largest retailers such as Walmart, Instacart, and Amazon Fresh, the success of these programs in affecting the usage and expenditure of SNAP recipients could provide useful insights to government and retailers.

This investigation expands on previous research into consumer adoption of online grocery shopping and the impact of consumer behavioral traits on the usage of online grocery shopping and purchasing decisions within online grocery markets. This study aims to contribute to the existing literature by identifying specific strategies that retailers can implement to overcome consumer barriers to online grocery shopping, ultimately increasing the adoption and usage of these services.

Objective Statement

The objectives of this research are to understand consumer attitudes and preferences toward online grocery shopping, identify groups of consumers within the population, and understand these clusters' demographics to allow for targeted marketing. Using these clusters and household information, we aim to understand how these preferences and qualities impact consumer willingness to purchase from online grocery stores and their expenditure habits.

Hypotheses

Based on our outlined objectives, our hypotheses follow in a similar format.

- Hypothesis 1: Consumer demographics will influence their decision to use online grocery services.
- Hypothesis 2: Consumer demographics will influence their in-person and online typical grocery expenditure.

- Hypothesis 3: Customers will be categorized into groups of individuals who care differently about temporal, financial, and trust aspects of online grocery shopping.
- Hypothesis 4: Consumer's cluster assignment will affect their decision to use online grocery services as well as their online and in-person typical grocery trip expenditure.

Chapter 2

Literature Review

Online Grocery Shopping Services

Online grocery stores and retailers refer to businesses that allow customers to purchase groceries and household items from their website or mobile application, with delivery options ranging from pickup to home delivery. These retailers fall under the umbrella of business-to-consumer (B2C) e-commerce, a section of e-commerce focused on sales between internet transactions of goods and services between a business and a private consumer (Statista 2023). Online grocery stores and retailers have gained popularity due to several factors, including convenience, time-saving benefits, and reduced impulse purchases (Statista 2023). Online grocery shopping (OGS) provides consumers a manner to obtain their necessary goods from a grocery store conveniently from their digital devices quickly and choose between picking up their order in person or having a delivery driver bring it to a designated location.

Home Delivery Services

With most retailers offering home delivery services or the option for curbside pickup, consumers can now enjoy the convenience of having their groceries delivered right to their doorstep. During a recent investigation, it was found that 60% of respondents use delivery while 40% pick up their online orders (Etumnu, 2020). Almost all U.S. online grocery stores and retailers offer same-day delivery for their products, but several offer rapid deliveries that arrive at a consumer's door within two hours or less of completing an order. Consumers value efficient delivery systems so online retailers also need to ensure that they provide an efficient delivery service to their consumers (Ganapathi, 2015). This theory should hold to our results and show that consumers value rapid delivery systems, but other results would lead to alternate conclusions and managerial implications. Consumers are not sensitive to minimum order requirements, but

they are found to be sensitive, in Bauerova's 2015 investigation of consumer behavior, to delivery times and prices. The most valuable quality of delivery systems to consumers was that delivery windows were accurate and convenient for the consumer (Bauerova, 2018).

Effective and reliable home delivery services can attract a wide range of customers, including those with limited mobility or access to transportation. Home delivery systems too possess the capacity to enable healthy eating. In a behavioral nutrition study, participants were subjected to the same diet program, but those who were forced to use online grocery services to purchase their groceries experienced significant weight loss compared to participants who shopped in physical stores. These participants increased their household food environments and decreased their consumption of impulse purchases (Gorin et. al. 2007). Therefore, home delivery services not only provide convenience but also have the potential to promote healthier eating habits and reduce impulsive purchases.

Convenience and Time-Saving

The primary attraction of online grocery stores and retailers is their convenience and time-saving factors, for customers whose schedules are too busy to allocate time for grocery shopping or who live in areas that do not have convenient access to brick-and-mortar supermarkets (Hand, 2008; de Magalhaes et. al., 2021). Hanus (2016) found that the most important advantages of online shopping are convenience and time savings. Similarly, Magalhaes (2021) found that consumers preferred the methods of shopping they perceived as requiring less time (traditional or online). Households with working adults and/or children have been found more likely to use online grocery stores (Pitts et. al., 2018), suggesting that online grocery stores and retailers serve as a practical solution for those who prefer the ease of shopping

for groceries from their homes. Incentivizing those who would not otherwise use online grocery retailers may require retailers to design marketing emphasizing the convenience and time-saving benefits of online grocery shopping (Shah and Tiwari, 2021).

Perceived Risk

Perceptions of online grocery shopping can often revolve around concerns regarding the quality and freshness of products, the usage of delivery, and the feasibility of shopping for one's weekly groceries using an app or website on an electronic device. A two-stage cluster analysis found that risk perception and trust of the retailer are factors that affect consumer perspective toward online grocery shopping but that the effect of the factors differentiates between age, income levels, and education (Brand et. al., 2020). This risk perception towards the feasibility of using online grocery shopping can be affected by producer action, as consumer risk is found to decrease with repeated usage of online grocery shopping (Mortimer et. al., 2016). This suggests that retailers need to encourage more adoption and frequent usage of online grocery shopping to reduce the perception of risk in using the service in general. A cluster analysis found that a segment of respondents was concerned with the chain of the supermarket and the online shops' trust (Frank et al., 2020). This study also found that the perceived social norm of the industry as a whole and the perceived general usage of a specific retailer aid in making consumers feel as if they can trust the process of online grocery shopping more (Frank et. al., 2020).

Mortimer's suggestion to address the issue of risk perception is to determine exactly which factors are causing consumers to perceive risk, such as the delivery processes and the quality of perishables (Mortimer et. al., 2016). In the same paper, Mortimer (2016) found that consumers who perceived higher risk were less likely to purchase perishable items, such as meat

and dairy, online. They were also less likely to use home delivery services, preferring to pick up their groceries in-store to ensure the quality and freshness of their products (Mortimer et. al., 2016, Hassan et. al., 2016). This secondary outlook proves especially insightful for observing consumer risk perception of perishable items as well as risk perception with the delivery process, both of which can be investigated through the observation of consumer attitudes towards these attributes.

Demographic Factors

Demographic factors such as age, income, regionality, and technology literacy can be significant barriers to the adoption and usage of online grocery stores and retailers (Bezirgani et. al., 2021). Income can be an impactful factor in the consumption of online grocery services due to the associated costs, inability to pay for the internet, and lack of access to digital devices. A relevant investigation found that those below the income threshold to receive SNAP benefits were hesitant to use online grocery shopping due to the high costs of delivery and tips but found that the desire to use the services exists and can occur when the associated costs are subsidized (Applehans et. al., 2013). A response to this has been reverse subsidization of online grocery shopping by allowing the usage of EBT, which when studied, has not been effective at driving usage of online grocery services and it has been suggested by investigators to encourage the USDA to invest in mechanisms to enhance consumer control and trust of perishable consumption online and to enhance the social norm of using online grocery services (Martinez et. al., 2018).

Age and technological barriers go hand in hand, as while older folks struggle and have very low current usage of online grocery shopping (4%) they show a high desire to use these services (Bezirgani et. al., 2018). It follows that theoretically, these older folks have a lower

percentage of their lives where they have had access to technology, especially the capacity to do their weekly shopping from their phones with the push of non-existent buttons. This technological barrier is not so significant as the perceived technological barrier to consumption of these services, which is found to be eliminated once the first usage has been completed (Bezirgani et. al., 2021). Although this demographic uses online grocery shopping relatively little, due to mobility issues, they may benefit from the convenience of online grocery shopping (Bezirgani, 2021). It has been suggested that investigations begin to observe the effect of geographical regionality on consumer perception and attitudes toward online grocery shopping, so in this investigation, we intend to build upon this suggestion (Bauerova, 2018).

Chapter 3

Materials and Methods

Data Collection

The data for this investigation was collected via an online Qualtrics survey from May to July 2023 across the US. The online survey was administrated using the online survey platform Qualtrics and distributed to the national representative consumer panels via online invitations. The survey was approved by the Institutional Review Board (IRB) at the University of Tennessee, Knoxville, before it was sent out. The final survey sample includes 1,530 respondents across the nation. Only primary household grocery shoppers, who were older than 18 years old, were qualified to participate in the survey.

Survey Design

The first component of the survey consisted of five parts, focusing on participants' basic demographics, then participants' knowledge and perception of online grocery shopping, and participants' internal reference price for a typical online and in-person grocery shopping trip. If participants did not shop online grocery shopping before, a set of barrier questions will be asked while for those who did shop online before, their attitudes towards online grocery shopping were examined. This set was followed by demographic questions designed to determine income, poverty status, employment status, household status and characteristics, vehicle ownership, and zip code. Demographic questions such as race, income, education, and employment were encoded as typically seen in the literature. To encode the vehicle ownership, respondents were given multiple options to describe their ownership but only if they answered that they owned a car personally they were encoded as owning a car and all others were encoded as not owning a car. Similarly, we asked respondents if they received Supplementary Nutrition Assistance Program benefits (SNAP), the Women, Infant, and Children Nutrition Program (WIC), or

neither. If they responded that they did receive SNAP, they were encoded as such with all others encoded as not receiving SNAP benefits. Finally, respondents had a selection of answers to choose their distance from their nearest grocery store, which were subjugated into three categories: those that live 0-5 miles from a grocery store, those that live 5-15 miles from their nearest grocery store, and those that live 15 or more miles from their nearest grocery store. These demographic factors will allow our findings to be viewed through the lens of any demographic factor, allowing for targeted recommendations for marketing to different clusters of consumers.

To find a way for online grocery stores to best target their consumers we approach this issue in a multi-stage form. The first stage we encountered was to use a series of nine questions that inquired about the attitudes of our survey respondents towards online grocery shopping. The first three questions in this questionnaire focused on the time-saving attributes that are hallmarked in the literature of online grocery shopping. The next set of questions focused on the financial cost associated with shopping and online grocery stores, followed by the last three questions, which focused on trust-related issues within an online grocery store. This set of nine questions allowed us to better understand how our respondents felt about shopping online and let us identify certain barriers or benefits that our respondents experienced in their grocery shopping history.

Data Analysis Method

Using these nine questions, we completed a K means cluster analysis to determine the clustering within our 1,530 respondents regarding their attitudes toward online grocery shopping. While K-means clustering is typically used for ordinal, variable clustering assignments, we chose to use K-means clustering in this scenario, due to experience with using the system for this type

of survey and past usage and literature that affirmed this methodology would be sufficient for determining robust results (Yenerall et. al., 2020). To begin the cluster analysis, we took the nine variable questions related to the attitudes of our respondent's online grocery shopping and ran means clustering on these variable sets. We began by running cluster analysis on the variables using different size clusters between two and five clusters. Using the elbow test, we identified that the ideal number of clusters was between three and four clusters. Following the elbow test, we use the silhouette between a three-cluster and a four-cluster analysis to determine that the silhouette was an analysis that determines the within clusters some of the squares and between clusters of squares, identifying the uniqueness of clusters. This allowed us to choose the cluster set that provided the most insight into the four cluster sets. From the silhouette analysis above, we can see the average silhouette is 0.15, with a peak of 0.17 and a minimum silhouette score of 0.09. As a result, we concluded that a four-cluster scenario is a more optimal case, and we verified it by comparing and investigating the loadings between the three-cluster scenario and the four-cluster scenario.

After we had finished digging through the cluster analysis and finalizing our ideal number of clusters, we wanted to further understand the demographic and so economic variables that were tied to our clusters. We completed a summary statistics table containing everything from an extensive breakdown of income brackets to all of the independent variables that were used in the regression. By using ANOVA testing to determine if there are statistically significant differences in these summarized variables between the clusters, we were able to understand what significant economic, demographic, and household differences persisted between our clusters.

Wanting to understand the impact of the clusters that we have identified on the online grocery shopping usage and expenditure of our respondents to simulate different clusters of

consumers within the United States, we utilized the logit model to analyze the impacts of our clusters and different personal characteristics of our respondents since our variable, indicating whether or not individuals have purchased from an online grocery store in the last six months, is a binary variable. Using the cluster groupings, allowed us to understand the impact of these clusters, socioeconomic conditions, household characteristics, and demographics on consumers' willingness to purchase online groceries in the last six months. We then move forward estimating, using the same variables, the impact that they have on the typical expenditure of our respondents within both traditional and online grocery stores by employing the ordinary least squares model. These results allow us to support the marketing of online grocery services to targeted groups of consumers, depending on their current attitudes towards online grocery shopping and the personal characteristics that they possess.

Model Specification

Our analysis consists of regression analysis, using the clusters that we identified in the first section, which aims to build a model that will predict the usage of online grocery stores in the last six months (a binary variable indicating yes or no), the amount spent on a typical in-person trip to the grocery store, and the amount spent on a typical purchase from an online grocery store. We used a logistic regression and two ordinary least square regressions in this study. Here are the model specifications:

Logistic Model:

$$\begin{aligned} \text{Usage of Online Grocery Shopping in Last 6 Months (Yes or No)} = & \beta_0 + \beta_1 * Age_i + \beta_2 \\ & * Income_i + \beta_3 * White_i + \beta_4 * Black_i + \beta_5 * Latinx_i + \beta_6 * College_i + \beta_7 * Grad_i + \beta_8 * \\ & Employed_i + \beta_9 * Car_i + \beta_{10} * Distance5to15_i + \beta_{11} * Distance15Plus_i + \beta_{12} * Children_i \\ & + \beta_{13} * SNAP_i + \beta_{14} * C1_i + \beta_{15} * C2_i + \beta_{16} * C4_i + \varepsilon_i \end{aligned}$$

Linear Model 1:

$$\begin{aligned} \text{TypicalInPersonGroceryShoppingSpending} = & \beta_0 + \beta_1 * \text{Age}_i + \beta_2 * \text{Income}_i + \beta_3 \\ & * \text{White}_i + \beta_4 * \text{Black}_i + \beta_5 * \text{Latinx}_i + \beta_6 * \text{College}_i + \beta_7 * \text{Grad}_i + \beta_8 \\ & * \text{Employed}_i + \beta_9 * \text{Car}_i + \beta_{10} * \text{Distance5to15}_i + \beta_{11} * \text{Distance15Plus}_i \\ & + \beta_{12} * \text{Children}_i + \beta_{13} * \text{SNAP}_i + \beta_{14} * C1_i + \beta_{15} * C2_i + \beta_{16} * C4_i + \varepsilon_i \end{aligned}$$

Linear Model 2:

$$\begin{aligned} \text{TypicalOnlineGroceryShoppingSpending} = & \beta_0 + \beta_1 * \text{Age}_i + \beta_2 * \text{Income}_i + \beta_3 * \text{White}_i \\ & + \beta_4 * \text{Black}_i + \beta_5 * \text{Latinx}_i + \beta_6 * \text{College}_i + \beta_7 * \text{Grad}_i + \beta_8 * \text{Employed}_i \\ & + \beta_9 * \text{Car}_i + \beta_{10} * \text{Distance5to15}_i + \beta_{11} * \text{Distance15Plus}_i + \beta_{12} \\ & * \text{Children}_i + \beta_{13} * \text{SNAP}_i + \beta_{14} * C1_i + \beta_{15} * C2_i + \beta_{16} * C4_i + \varepsilon_i \end{aligned}$$

We included not only the clusters that we identified in the beginning, using the indifferent cluster as our base, but we, too included major demographic information such as age and race, so economic information, such as income, car ownership, and household information from the existence of children within the household to the distance of the household from their nearest grocery store for each individual i . In this regression, $C1_i$, $C2_i$, and $C4_i$ stand for the dummy variables indicating the cluster group assignment for individual i . The third cluster group (the indifferent cluster, $C3_i$) was used as the baseline benchmark. Based on past literature, we used logistic regression to analyze the dependent variable of whether or not somebody has purchased from an online grocery store in the last six months. Finally, we used ordinary least squares linear regression analysis to analyze the impact of our independent variables on our expenditure variables.

Chapter 4

Results and Discussion

Cluster Analysis

To understand the clustering of consumers' attitudes towards the financial, temporal, and trust-related attributes of online grocery shopping, we followed a plan designed to allow for robust results. We began our investigation by using the nine attitude questions mentioned earlier to cluster our respondents into the ideal number of clusters and use the silhouette and elbow methods to determine this ideal number. We then followed by running summary statistics on the four clusters to understand the demographic, socioeconomic, and household characteristics of each cluster. To begin, using the elbow test and the silhouette method, we finalized a design for four cluster groups (Figure 4.1). As an extra robustness check, we investigate the loadings between the third and the fourth cluster completely. Within the third cluster, we did not find that the analysis was as robust as the fourth cluster, as there was an online grocery shopping supporter group, a hater group, and a large indifferent group. When we dug further into the four cluster means analysis, we discovered that much of the larger and different group was unrepresented. In this four-cluster analysis, the loadings were much more descriptive. The fourth cluster was identified as a trade-off group, individuals who both saw the value of online grocery shopping and saw that there were barriers to their usage of online grocery shopping.

Once we have determined that four clusters are the ideal number of clusters, we see the first breakdown of our respondents into these clusters as we too, confirm that our hypothesis three is correct, respondents can be divided amongst the three factors of temporal, financial, and trust aspects of shopping for groceries online. The results of the summary statistics are summarized in Table 4.2, yet Table 4.1 represents the cluster loadings of the four clusters. The smallest cluster is our online grocery shopping lover cluster, with 15.4% of our sample falling in

this group (Table 4.1). This group is identified as those who see mostly the positive value in shopping for groceries online and, at worst, hold indifferent opinions towards online grocery shopping attributes. More than a third of our sample fell into the largest cluster identified, with 37.1% of respondents falling into the trade-off group, which are respondents who both see the value and the potential barriers to shopping for groceries online (Table 4.1). The second largest group is our indifference cluster, in which respondents remained around neither a degree nor disagree loading for all but one attribute. This group was comprised of 21.7% of our sample. The third largest and final group was the online grocery shopping haters, who comprise 25.8% of our sample and represent those who experienced agreement with questions related to barriers towards online grocery shopping and who did not identify value within online grocery shopping (Table 4.1).

Specifically, the OGS lover cluster found they were different regarding the time that they spent shopping but disagreed that finding time to have online groceries delivered is hard (Table 4.1). They disagree that buying groceries online doesn't save money, indicating that they still see the value in online grocery shopping as a way to save time and money while getting their groceries. The lover cluster was indifferent regarding whether online grocery delivery fees were too high, and they perceived that buying online groceries does save money and that stores don't charge more for the products that they sell online. Finally, this cluster is indifferent to whether someone else selects their food and is not concerned with sharing their personal information online. This cluster group values the time-related attributes and does not think that extra charge is an issue. The cluster group also shows a tendency to trust the services provided by OGS (Table 4.1), finalizing that our first cluster follows our hypothesis three.

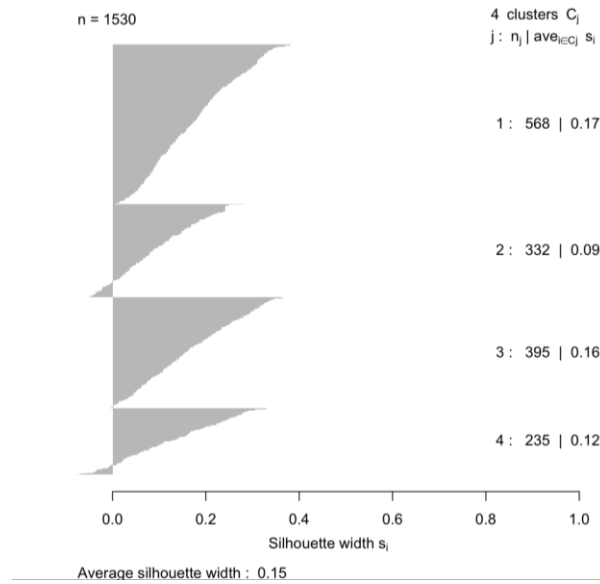


Figure 4.1. Silhouette Plot of 4 Cluster K-Means Analysis.

Table 4.1. Table of the Four Cluster Analysis.

Attitude Questionnaire Cluster Analysis				
Variable	OGS Lover Cluster	Tradeoff Cluster	Indifference Cluster	OGS Hater Cluster
Size with Total Sample of 1,530 Respondents	235 (15.4%)	568 (37.1%)	332 (21.7%)	395 (25.8%)
Silhouette Width	0.12	0.17	0.09	0.16
Within Cluster Sum of Squares	2553.12	4631.34	3558.41	3443.67
Time Related Attitude Statements				
I don't like to spend lot of time shopping.	2.99	2.13	3.33	2.78
Finding a time to have online groceries delivered is hard.	1.38	1.79	3.14	3.22
Buying groceries online doesn't save much time.	1.28	1.67	2.88	3.52
Spending Related Attitude Statements				
Delivery fees for online groceries are too high.	2.65	4.22	3.13	4.47
Buying groceries online doesn't save money.	1.91	2.65	3.52	4.29
Stores charge more for groceries they sell online.	2.42	3.77	3.04	3.89
Trust Related Attitude Statements				
I don't trust someone else to select my food.	2.67	4.13	3.11	4.59
It's risky sharing personal information online.	2.75	4.09	2.64	4.07
It's risky having someone else pick out my fruits and vegetables.	2.09	3.79	3.29	4.66
Decision Criteria: If cluster loading greater than 3.5 or less than 2.5, loading encoded as agreement or disagreement, respectively.	I agree with this statement.		I disagree with this statement.	

The trade-off cluster is where we see the variation in people's attitudes towards online grocery shopping within the cluster, as we see that these respondents don't mind spending a lot of time shopping for groceries and they don't think that finding an online delivery slot is difficult (Table 4.1). While they think that buying groceries online does save time, they have fears regarding the spending and trust attributes of online grocery shopping. The trade-off cluster perceives that delivery fees for online groceries are too high and that stores do charge more for the food that they sell online. Online grocery shopping saves money, but they do not trust someone else to select their food and feel as if it is risky for someone else to pick fruits and vegetables. Finally, they perceived that it was risky to share their information online. Therefore, this cluster group generally appreciated the time-saving aspects of OGS while being concerned about the extra charge and trust-related issues, which caused them to make tradeoffs when facing OGS (Table 4.1). These tradeoffs are exactly what was hypothesized in hypothesis three, that consumers would be divided amongst their perception of these factors.

The indifferent cluster maintains indifference across every single question except that they do agree that buying online groceries doesn't save money (Table 4.1). This indicates this cluster group recognizes the time-saving aspects of shopping online. Similarly to this opinion, the OGS hater cluster does agree that buying online groceries doesn't save money (Table 4.1). They are indifferent to whether they like to spend a lot of time shopping and the difficulty of finding a delivery timeslot for the groceries to be brought to them. They agree that buying online groceries doesn't save money and buying a new line grocery doesn't save time. They think that the delivery fees for online groceries are too high and that stores charge more for the groceries that they sell online. They don't trust someone else to select their food, and they perceive that having someone else pick out their fruits and vegetables is risky. Finally, this cluster perceived

that sharing their personal information online was risky. As a result, this cluster group was generally concerned about spending more when shopping for groceries online and about trust-related issues. This cluster group of respondents was not so likely to convert to OGS very easily unless significant changes were made (Table 4.1), but this does show that they are separated in their attitudes towards financial and trust related aspects of OGS finalizing the confirmation of hypothesis three.

Table 4.2 represents the breakdown of the summary statistics of the sample in our study and the four clusters identified. For the total sample, we have a count of 1530 individuals with an age of 48.64 (Table 4.2). 56.6% of our sample are employed and 80.33% are owners of a car. On average 32.94% have children that live in their household with them and 23.66% have a four-year degree, while 11.83% hold a graduate or postgraduate degree. In terms of distance from grocery stores, 62.22% of our sample lives within 5 miles of a grocery store and 30.78% of our sample lives 5 to 15 miles away from a grocery store, leaving 6.41% of our sample living more than 15 miles from a grocery store. The average amount that our total sample spent on a typical in-person shopping trip to the grocery store was \$133.3 and the average typical amount that was spent online grocery shopping trip was \$125. Our sample has 56.73% of the respondents have used online in the last six months, which is significantly higher than found in the literature, which is reasonable since we only collected the data from online outlet (Table 4.2).

The income breakdown of this total sample indicates that there are 14.6% of respondents have a household income of zero to \$25,000, 20.9% of respondents have an income of \$25,000-\$49,999, 35.5% of our sample as an income between \$50,000 and \$99,999, 18.5% of our sample has an income of \$100,000-\$149,999, and 9% of our sample has a household income of \$150,000 or more (Table 4.2). The percentage of respondents identifying white is 69.65%, and

the percentage identifying as black or African American is 10.62%. Respondents who identify as Hispanic or Latin account for 13% of our sample, 4.1% of our sample identifies as Asian, and 2% of our sample identify as another race. Finally, 20.13% of our sample identifies that they receive SNAP benefits (Table 4.2).

The first cluster is the online grocery store lover cluster, which makes up 235 respondents at 15.4% of the total sample size, the smallest cluster (Table 4.2). The average age is the same as the respondents aged 48.6 years old, and the second highest employment at 59.57%. This cluster has the lowest amount of people that own a car at 66.4% and the highest amount of children in the household at 38.3% amongst the clusters. This is the cluster with the highest amount of education, with 27.2% of people possessing a four-year degree and 11.9% of the people possessing a post-graduate degree. The least amount of people within this cluster live within 5 miles of the nearest grocery store and the highest amount of people that live between 5 and 15 miles from the grocery store at 38.7%, followed by 5.5% of the population living more than 15 miles away from the grocery store, which is the lowest amongst the clusters. This group spends the second highest amount on their typical in-person grocery shopping trip at \$134.3 per trip and the highest amount on their online grocery shopping orders at \$139.7 on average. This cluster has the astoundingly highest amount of experience using online grocery shopping with 92.7% of the cluster having experience shopping online for groceries in the last six months. In the lowest income, 14.47% of this cluster exists, while 21.3% of the cluster lives within the second income bracket. 36.17% of the cluster falls within the third income bracket, and 18% of the sample falls within the fourth highest income bracket. Those who receive snap benefits comprise 21% of this cluster (Table 4.2).

Table 4.2 Table of Summary Statistics of Relevant Variables for Regression

Variable	Total Sample	OGS Lover Cluster	Tradeoff Cluster	Indifferent Cluster	OGS Hater Cluster	ANOVA P-Value
Respondent Has Purchased Grocery Online in Last 6 Month	56.73%	92.77%	74.47%	46.99%	17.97%	6.98e-05 ***
Household Income of \$0 - \$24,999	14.64%	14.47%	13.03%	12.95%	18.48%	0.122
Household Income of \$25,000 - \$49,999	20.85%	21.28%	18.13%	19.88%	25.32%	0.0445 *
Household Income of \$50,000 - \$99,999	35.49%	36.17%	36.80%	39.16%	30.13%	0.217
Household Income of \$100,000 - \$149,999	18.50%	18.72%	19.37%	18.37%	17.22%	0.592
Household Income of \$150,000+	9.02%	8.94%	11.62%	7.23%	6.84%	0.0533 .
White	69.65%	65.26%	63.80%	70.76%	79.19%	0.0144 *
Black or African American	10.62%	16.43%	13.00%	7.97%	6.22%	0.784
Hispanic or Latinx	13.01%	13.62%	17.60%	13.29%	6.22%	0.000885 ***
Asian	4.12%	2.35%	3.60%	5.32%	4.86%	0.787
Other	2.09%	2.13%	1.58%	2.11%	2.78%	0.347
Receive SNAP	20.13%	21.12%	24.77%	20.19%	13.07%	0.00466 **

There are 568 respondents within the trade-off cluster or 37.1% of our sample (Table 4.2). The average age of this cluster is 43.85, lower than the average of our total samples. This cluster has a higher amount of employed individuals at 65.5% and 80.99% of these respondents on a car. A higher number of these respondents were children at 38.2% compared to our total sample. 23% of our respondents in this cluster hold a four-year degree, and 12.32% of our cluster has a graduate or postgraduate degree. The breakdown of this cluster's distance from the nearest grocery store is nearly identical to that of the total sample with the exception that an extra percent of individuals live 5 to 15 miles from a grocery store. This cluster spends more than the total sample on their typical in-person shopping trip at \$145.48 and they spend a little less on the typical online grocery shopping trip at \$123.7. Nearly 20% more of this cluster has purchased groceries in the last six months online at 74.47% and the income breakdown follows as such: 13% of people are in the lowest income bracket, 18% of people are in the second highest bracket, 36.8% of people have a household income between \$50,000 and \$99,000, 19.37% within the cluster have a household income in the fourth highest bracket, and 11.62% of people have a household income of greater than \$150,000 (Table 4.2). This cluster is more diverse than the sample as 63% of the respondents identify as white, 13% identify as black or African American, 17.6% of this cluster identify as Hispanic or Latin, 3.6% identify as Asian and 1.5% identify as another race. This cluster has a higher percentage of people who receive SNAP benefits at 24.77% (Table 4.2).

The indifference cluster comprises 332 people, which is 21.7% of our total sample (Table 4.2). The average age is near the total sample average at 47.73%. Just right near the total sample, 56.6% of this cluster is employed, and 82.5% of this cluster owns a car. Reflecting the total sample closely, 33.4% of people in this cluster have children within the household, 20.78%

hold a four-year degree, and 11.5% of people hold a graduate or post-graduate degree within the cluster. 64% of people within the cluster live under 5 miles from the grocery store, followed by 27.7% of people who live 5 to 15 miles from the grocery store, and 7% of people who live more than 15 miles from a grocery store. This cluster reflects the averages for the total sample nearly as this cluster average spends \$134.4 on their typical in-person shopping experience and \$121 on their online grocery shopping experience. Nearly half of the cluster has shopped online in the last six months for groceries at 46.99%. In the lowest income bracket, there are 12.95% of this cluster, followed by 19.88% in the second highest bracket, 39.16% holding income between \$50,000 and \$99,999. 18.37% of this cluster has a household income between 100,000 and \$149,999 and 7.2% of this cluster have an income in their household greater than \$150,000 (Table 4.2). Those who identify as white comprise, 70.76% of this cluster followed by 8% who identifies black or African American, with 13.29% of respondents in the cluster identifying as Hispanic or Latin necks. 5.32% of this cluster identifies as Asian and 2% of this cluster identifies as another race, 21.9% of this cluster identifies as receiving SNAP (Table 4.2).

The final cluster is the online grocery shopping hater cluster, which is comprised of 395 individuals, which equates to 25.8% of our total sample with an average age older than all of the other clusters of 56.6 years old (Table 4.2). A lower percentage of these respondents are employed at only 42% and 85.8% of these individuals own a car. At the lowest percentage of children in the household, only 21.8% of these respondents have children in the household and 24% of these individuals hold a four-year degree and 11.4% of these individuals hold a graduate or post-graduate degree. Those that live under 5 miles from the grocery store make up 65.57% of the cluster all those that live between 5 and 15 miles from the nearest grocery store make up 26.8% of the cluster, followed by 6.58% who live more than 15 miles from the grocery store.

This cluster spends the least on both in-person shopping and typical online grocery shopping trips, with \$114.1 spent on the typical in-person, shopping trip and \$96.6 on the typical online grocery trip. This cluster has the lowest amount of experience with online grocery shopping, as only 17.97% of this cluster has purchased online grocery in the last six months. This cluster has the highest amount of respondents within the \$0 to \$24,999 income bracket and the highest amount within the \$25,000-\$49,999 range. It has the lowest percentage of people within the next three income brackets, with 30.1% of the population within the third income bracket, 17.2% of the population within the fourth highest income bracket, and 6.8% of the population within the highest income bracket (Table 4.2). This cluster has the highest amount of people that identifies as white with 79% followed by 6.2% of the population identifying as black or African American and 6.2% of the population, identifying as Hispanic or Latino. Exactly 4.86% of the sample identifies as Asian 2.7% of the population identifies as other, with the lowest percentage of individuals that receive SNAP being in this cluster at 13.07%.

We too, ran the ANOVA test to gain the P values that determine if there are differences between the four clusters that are statistically significant. We found that there is a statistically significant difference between the clusters for the respondent's age, employment status, car ownership, and whether children are in the household (Table 4.2). We also found that there's a statistically significant difference in the typical in person, shopping expenditure between the clusters, and a very significant difference between the clusters on whether or not the respondent has purchased groceries within the last six months online. We find a statistically significant difference between the clusters for the percentage of individuals with a household income between \$25,000 and \$49,999 and for households with an income greater than \$150,000 (Table 4.2). A statistically significant difference exists between two of the racial categories,

those that identify as white, and those that identify as Hispanic or Latinx. Finally, we find a statistically significant difference between the percentage of houses that receive SNAP benefits (Table 4.2).

Regression Analysis

In Table 4.3, we present the regression results of the estimation of willingness to purchase groceries online and the related expenditure factors mentioned in the introduction of the section. We present not only the loadings of the regression results, but also the count of the dependent variable being analyzed, and the log-likelihood, in the case of the logic model, and the R squared value for our ordinary squares estimations. First, in the logistic regression where we examined the respondent's behavior of whether to purchase groceries online in the last six months, we replaced the regression coefficients with the average marginal effect. We found that eight variables have a significant impact on the online grocery shopping usage of our respondents. With 1341 individuals being analyzed in this regression, we find that as the age increases one unit, this decreases the likelihood of the respondent purchasing groceries online by 0.2% (Table 4.3). We find that if the individual is Hispanic or Latinx, this increases the likelihood of purchasing groceries online in the last six months by 10.74%, followed by the case of children existing within the home if they do exist the household likelihood of purchasing groceries online in the last six months increases by 11.73%. If the individual is employed, this leads to a 7.28% increase in the likelihood that the individual would purchase groceries online in the last six months and if the household receives SNAP benefits, this increases the likelihood that the house will purchase groceries online in the last six months by 6.43%. These results confirm that the demographics of our respondents do affect their usage of online grocery shopping in the last six months, supporting our first hypothesis. Looking into the clusters we find that if the

individual is within the trade-off cluster, this increases the likelihood of the individual purchasing groceries online in the last six months by 19.04% when compared with the indifferent cluster supporting our fourth hypothesis that cluster assignment impacts the usage of OGS. Using the indifferent cluster as the baseline, we see that the individuals who went into the online grocery shopping hater cluster decreased the likelihood that they would purchase groceries online in the last six months by 17.88%. Finally, when compared with the indifferent cluster, we find that if the individuals within the online grocery shopping lover cluster this increases the likelihood that the individual has purchased groceries online in the last six months by 48.43% (Table 4.3). These three significant results support our fourth hypothesis that cluster assignment impacts consumer's usage of online grocery shopping services in the last six months.

With a count of 1,325 respondents, we analyze the impact of these variables on a typical in-person expenditure for our respondents using the OLS (Table 4.3). We find that if the individual increases one year in age, this decreases the expenditure by \$0.58. If the individual increases one income bracket from the nine provided to them as options, this increases the typical in-person spending by \$3.56 and if children live in the household, we see that this increases the typical in-person spending by \$24.8. If the individual holds a graduate or post-graduate degree, this decreases their typical in-person spending by \$17.74. If the individual owns a car, this increases their expenditure by \$22.8 as opposed to not owning a car. Compared to individuals who live within 5 miles of the nearest grocery store, if they live between 5 and 15 miles from a grocery store this increases their expenditure for a typical in-person trip to the grocery store by \$19.57. If individuals live more than 15 miles from a grocery store, this increases their expenditure by \$39.95.

Table 4.3 Results of Logit and OLS Regressions

Regression Results of Estimation of Willingness to Purchase Groceries Online and Expenditure.			
Dependent Variable and Model	Purchase of Online Grocery in Past 6 Month (Logit)	Typical In Person Spending (OLS)	Typical Online Spending (OLS)
Independent Variable	Estimate	Estimate	Estimate
Intercept	-0.4286	121.7260 ***	90.8384 ***
Respondent Age	-0.0020 **	-0.5827 **	-0.3316
Household Income Before Taxes	-0.0002	3.5588 **	3.5233 *
White	0.0596	-16.0854	3.6392
Black or African American	0.0700	7.5088	3.6823
Hispanic or Latinx	0.1074 *	19.5121	14.5663
Children Exist In Household	0.1173 ***	24.8016 ***	14.3308 *
4 Year Degree Holder	0.0224	-9.0586	-8.9455
Graduate or Postgraduate Degree Holder	0.0228	-17.7387 *	-11.1529
Employed	0.0728 **	-3.1662	13.1618
Owns Car	-0.0459	22.8278 ***	0.1383
Lives 5-15 Miles from Grocery Store	-0.0049	19.5740 ***	25.7619 ***
Lives More than 15 Miles From Grocery Store	-0.0632	39.9467 ***	16.4454
Household Receives SNAP	0.0643 *	8.9983	5.2385
Tradeoff Cluster	0.1904 ***	-0.4163	-3.4094
Online Grocery Shopping Hater Cluster	-0.1788 ***	-14.9850 *	-30.3368 *
Online Grocery Shopping Lover Cluster	0.4843 ***	-5.2035	12.6615
Count of Dependent Variable	1,341	1,325	715

Table 4.3 Continued

Log-Likelihood or R-Squared	-628.0599	0.1213	0.0649
------------------------------------	-----------	--------	--------

These significant demographic factors of our respondents support our second hypothesis that consumer demographics impact their expenditure on in-person grocery shopping. None of the clusters are significant except that if an individual is within the online grocery shopping hater cluster, this decreases their expenditure compared to the indifferent cluster by \$14.98 for a typical in-person grocery shopping trip (Table 4.3) still demonstrating our hypothesis four holds solidly. Finally, 715 respondents were used to analyze their typical online grocery shopping expenditure using OLS (Table 4.3). If the individual increases one of the nine income brackets, this increases the expenditure by \$3.52. If the individual has children that live within their household, this increases their expenditure by \$14.33. An individual lives 5 to 15 miles from the grocery store as opposed to living within 5 miles of the grocery store, which increases their online purchase expenditure by \$25.76 and confirms our second hypothesis. Finally, if the individual is within the online grocery shopping hater cluster as compared to the indifferent cluster. This decreases their expenditure online for a typical grocery shopping trip by \$30.34 (Table 4.3) while too confirming our fourth hypothesis that cluster assignment does impact consumer usage of OGS, in-person shopping expenditure, and OGS expenditure.

Robustness Check

To ensure that our models were as robust as capable, we ran variance inflation factor testing on all three regressions to account for multicollinearity within the models. Additionally, to ensure that the variable accounting for whether the individual household receives SNAP didn't negatively impact the coefficient estimation of the income variable, we dropped the SNAP variable from the model. We found that the most optimal log-likelihood score possible is -628.06 for our logic model, and for our typical in-person spending expenditure estimation, we found that the R-squared that was optimal was 0.1213 (Table 4.3). Our optimal R-squared value for the

typical online grocery order spending estimation is 0.065. From Table 4.4, we see that all variance inflation factor scores for every variable for each model are under the threshold of 5, indicating that the variables themselves do not suffer from multicollinearity. Now that we have checked our model statistics and for multicollinearity, we must investigate whether the SNAP variable skewed our estimations.

From Table 4.5, in comparison with Table 4.3, we find very similar results when we drop the SNAP variable. In the case of our estimation of the usage of online groceries in the past six months, we find that age, identification is Hispanic, or Latinx, children existing in the household, whether an individual is employed or not, and the clusters are still significant, and the coefficient resembles closely those found in the initially presented regression results. We find that as we continue estimating typical in-person and online grocery shopping expenditures, very similar results are found, indicating that including the SNAP variable into the equation does not hurt the results by skewing the estimated coefficients. This robustness check was implemented to ensure that this variable was not statistically significant because of multicollinearity or another misspecification error, and we were unable to locate any misspecification errors despite searching for them. Therefore, we decided that the original regression results presented in Table 4.3 are robust, as the final results.

Table 4.4. VIF test for the regressions to test multicollinearity.

Variance Inflation Factor Scores of Regression Results Estimation of Willingness to Purchase Groceries Online and Expenditure.			
Independent Variable	Purchase of Online Grocery in Past 6 Month (Logit)	Typical In Person Spending (OLS)	Typical Online Spending (OLS)
	Estimate	Estimate	Estimate
Respondent Age	1.6866	1.8506	1.6971
Household Income Before Taxes	1.2661	1.2967	1.3277
White	3.4465	3.6748	4.5745
Black or African American	2.2213	2.3888	2.9058
Hispanic or Latinx	2.3998	2.6265	3.2911
Children Exist In Household	1.2166	1.2771	1.2175
4 Year Degree Holder	1.0902	1.0981	1.1073
Graduate or Postgraduate Degree Holder	1.1169	1.1189	1.1287
Employed	1.4192	1.5403	1.5095
Owns Car	1.1239	1.1431	1.1384
Lives 5-15 Miles from Grocery Store	1.0443	1.0501	1.0539
Lives More than 15 Miles From Grocery Store	1.0483	1.0438	1.05697
Household Receives SNAP	1.1204	1.1450	1.1431
Tradeoff Cluster	1.4093	1.7397	2.0650
Online Grocery Shopping Hater Cluster	1.3591	1.7070	1.3897
Online Grocery Shopping Supporter Cluster	1.1477	1.5206	1.9952

Table 4.5. Regression for Robustness Check.

Regression Results Estimation of Willingness to Purchase Groceries Online and Expenditure Without SNAP Variable			
Dependent Variable and Model	Usage of Online Grocery in Past 6 Month (Logit)	Typical In Person Spending (OLS)	Typical Online Spending (OLS)
Independent Variable	Estimate	Estimate	Estimate
Intercept	-0.0557	124.8708 ***	97.5921 ***
Respondent Age	-0.0148 **	-0.5688 **	-0.3470
Household Income Before Taxes	-0.0252	2.7595 *	2.2615
White	0.3836	-16.0328	3.4774
Black or African American	0.4062	11.7758	4.1161
Hispanic or Latinx	0.6604 *	22.7785 .	13.1502
Children Exist In Household	0.7757 ***	28.5127 ***	16.2324 *
4 Year Degree Holder	0.1199	-9.5359	-9.1599
Graduate or Postgraduate Degree Holder	0.1527	-17.1754 *	-9.3111
Employed	0.5036 **	-2.3047	13.3693
Owns Car	-0.3282 .	22.5419 ***	3.6207
Lives 5-15 Miles from Grocery Store	-0.0636	18.2187 **	25.2162 ***
Lives More than 15 Miles From Grocery Store	-0.3701	40.5997 ***	18.6241
Tradeoff Cluster	1.2684 ***	1.0353	-4.7978
Online Grocery Shopping Hater Cluster	-1.1912 ***	-13.4637 .	-27.8638 *
Online Grocery Shopping Hater Cluster	3.0934 ***	-3.5352	12.3888
Count of Dependent Variable	1,341	1,325	715

Table 4.5 Continued

Log-Likelihood or R-Squared	-648.6629	0.1232	0.06108
------------------------------------	-----------	--------	---------

Chapter 5

Conclusions and Recommendations

Implications

In this investigation, we have aimed to create a clustering of consumers who have and don't have experience in online grocery shopping so that we may support the targeted marketing of these online grocery services both to incentivize the usage and expenditure within these markets. We created this clustering and developed clusters that surround the attitudes of respondents towards financial temporal and trust-related attributes for online grocery shopping. We decided that the ideal number of clusters was four clusters and it provided us insight into four segments of consumers: those who love online grocery shopping, those who see the value and see barriers towards online grocery shopping, those who are different about online grocery shopping, and finally those who hate online grocery shopping. Cluster analysis is the most robust, as it is, the most unique according to a silhouette score investigation, and it rests on the elbow of uniqueness by using the elbow test.

Within the first and smallest cluster, the online grocery shopping lover cluster, they are not concerned whether or not somebody selects their food and are indifferent to sharing their personal information online. This cluster also disagrees that it's risky to have somebody else pick their fruits and vegetables, likely due to their high level of usage of the services. This group of people has the most experience shopping online, with 92% of the cluster having had experience shopping online for groceries in the last six months. This cluster closely resembles the age of the total sample at 48.6 years old, and they do have the highest employment rate nearly 60% employed they have the lowest number of people that own a car within this cluster, which would support their usage as they don't all have the means to reach a grocery store by themselves and they also have the highest number of children in the household, which follows the theory that

more time-starved individuals use online grocery more. This cluster has the highest amount of people who have gained a degree between a four-year degree and a postgraduate degree, indicating that with education and the opportunities to use technology, these individuals can assess the value and have simple entry into this market. The few people within this cluster live within 5 miles of their nearest grocery store, and most live between 5 and 15 miles from the grocery store, indicating that the convenience of savings from fighting traffic to go to the grocery store incentivizes higher usage. Previous utilization in the literature has led to higher expenditure, and this track as this group spends the most amount of money for their typical online grocery order. This group is the most diverse with the lowest number of people who identify as white and those who receive SNAP benefits comprise 21% of the cluster. This cluster group is the foundation and includes major customers of today's online grocery shopping platforms. Since over 90% of the respondents in this group already used OGS, the online outlet could use this cluster group to study their shopping habits and patterns online to seek opportunities to expand their purchasing frequency and usage online.

Within the largest cluster, the trade-off cluster is where we see retailers have the highest opportunity to convert users to their service through target and marketing because they already see the value of online grocery shopping, yet they have certain barriers that must be addressed. This cluster doesn't mind spending a lot of time shopping for groceries, and they also don't see difficulty finding a timeslot for delivery as a barrier. They also believe that buying groceries online saves a lot of time, and while this cluster identifies that there are time-related benefits to shopping online, they fear that the delivery fees for grocery shopping online are too high. They are indifferent as to whether buying groceries online saves money, yet they believe that stores charge more for groceries they sell online. This cluster has issues with the trust associated with

shopping online for groceries, as they do not trust someone to select their food, and they do not trust somebody to pick out their fruits and vegetables for them while making an online order, while also being fearful of sharing their personal information online. This trade-off can be addressed through targeted marketing towards the demographic and household characteristics of the cluster to emphasize the related benefits of shopping online while working to dispel the spending and trust barriers that this cluster faces. This cluster is younger than the total sample, and they have a higher amount of employed individuals and children in their household compared with the total sample. This cluster represents the average total sample, education level, distance from their nearest grocery stores, and income distribution, indicating that this group is the general consumer, which is also supported by being the largest group in the cluster analysis. This group spends a little bit more on their typical in-person shopping trip, and they have one of the higher percentages of people that receive SNAP at 24.77%. Noticing that this cluster takes the largest composition of the total sample (over 37%), marketers find it great potential when addressing the concerns of those potential customers. The respondents in this cluster group already observed the benefits of the time-saving aspects of OGS while being concerned about the extra expenses and trust-related issues. If they were given a discount in terms of coupons, free delivery, or guaranteed product quality, those customers could be converted into the OGS lover group in the future.

The indifference cluster is the third largest group, and they maintain indifference amongst all perceptions towards online grocery shopping that they do not save money. The summary statistics for this cluster closely follow the total sample average specifically the age, employment, rate, and car ownership. Again, the distance from grocery stores, the typical expenditure, both online and in person, and the income and racial distributions follow along the

total sample averages, indicating that this group is the other segment of the typical consumer that online grocery stores would market if they are generally marketing, the most significant aspect of these consumers that is nearly 47% of them have shopped online for groceries in the last six months, indicating that they are not fully convinced about the benefits of online grocery and they perceived that from their experience shopping for groceries online can be more expensive. Compared with shopping for groceries in person, this perception tracks as there are almost always additional service or delivery fees related to shopping for groceries online, but this group does not identify that there are barriers to them using groceries online other than the cost. Currently, respondents in this cluster group were more neutral to the OGS. Unlike respondents in the tradeoff cluster, they are not typically attracted by any of the benefits brought by the OGS nor dislike any of the concerns. Since they don't hold the negative opinions about the OGS like the tradeoff cluster, they might be converted and turned into potential customers once they find anything positive about using OGS. Therefore, this cluster group can also be viewed as a potential customer group of OGS.

The last, and potentially the hardest to convince to adopt online grocery shopping is the online grocery shopping hater cluster, as this group sees negative throughout the three different attribute groups regarding online grocery shopping. They believe that the delivery fees for groceries online are too high and that stores charge more for the groceries they sell. They perceive that it is not trustworthy to allow someone to select their food or their produce, while the group is indifferent regarding whether they like to spend a lot of time shopping and the difficulty of finding a delivery timeslot for the groceries to be brought to them, which makes sense since they have the least amount of experience with services in the last six months of all the clusters. They also perceive that it is risky to share personal information online, which

supports the fact that this group finds the most amount of barriers within the online grocery store existence. These negative perceptions track as this group is the oldest amount of respondents at 56.6 years old on average, and they have the least amount of experience out of all the clusters with only 17.97%, both impactful attributes when it comes to whether or not somebody chooses to shop online and enjoys it from the literature. A lower percentage of these individuals are employed and only 42% and 85.8% of these individuals own a car, supporting the theory that these individuals are retired. They also have the lowest amount of children within the household, which would indicate that this group of individuals is likely an older group of people who do not see the necessity or benefit of shopping online for groceries. Based on having available time to go shop in person, this group of individuals spins the lease on their typical grocery shopping trip both in person and online, and they have the highest amount of respondents within the zero-dollar to \$24,999 income bracket in the highest amount of individuals amongst the clusters within the \$25,000 to the \$49,999 range of income further supporting the concept that these individuals are relying on pensions or independent retirement accounts. This cluster has the highest amount of people who identify as white in the lowest amount of individuals who receive SNAP coming in at only 13% of the cluster. This cluster, accounting for roughly one-fourth of the sample in the study, is probably the hardest group to consider as future OGS customers due to concerns about extra charges and trust-related issues. That said, this is probably the group that deserves the most attention if online outlets and service providers want to find places to improve their services online.

The literature defines age, education, income, and experience with shopping online as major indicators of an individual's online grocery shopping usage, which is very evident in our investigation. Age typically hurts the history of individual shopping online in the literature, and

this is also found in our investigation. Income is usually found to increase the usage of online grocery shopping within the market, and this resonates with our investigation. Those who have the most education are likely to have a higher level of experience with technology, and this has been found to have a positive impact on the usage of online grocery shopping. Usage in the past has been found to affect the attitudes that respondents in the literature hold towards online grocery shopping, and this resonates in our investigation, as the two most clusters (the online grocery shopping lovers and haters) have the highest and the lowest usage within the last six months, respectively.

The starkest finding that goes against the literature is resounding results regarding the impact of receiving SNAP on online grocery shopping, as it has always had a positive impact on the usage of online grocery shopping in the last six months of our investigation. In the literature, there have been many investigations on how to increase online grocery shopping by SNAP recipients by changing the way that individuals can use their SNAP benefits within these retailers, and our investigation differs from most, as we give us a six-month period for respondents to indicate whether they have shopped online for groceries or not whereas most investigations only use terminology indicates whether or not individuals have used online within the last month. Based on all of the research that occurred over the last five years toward increasing the usage of online grocery shopping by SNAP recipients, it is theoretically feasible that within the last six months, based on all of the adoption policies by major retailers that are results that indicate that being a SNAP recipient increases the likelihood of shopping for groceries online, which resonates with the fact that major OGS providers are encouraging and facilitating SNAP recipients to utilize their benefits online. This investigation's goal has been to support retailers' marketing strategies to optimize their efficacy in bringing new shoppers into

their market and incentivizing them to spend additional money per purchase. Through a detailed cluster analysis, summary statistic, breakdown, regression analysis, and ensuring robustness we have been able to provide insights into how a firm could incentivize different segments of consumers through online marketing to both enter their online grocery shopping market and increase their expenditure.

Limitations and Future Research

Despite the contributions of this study, there are some limitations and unanswered questions that could be solved by future research. First of all, this research only focused on the cluster groups of consumers and investigated the behavior of adopting OGS or not but did not study the behavior and tradeoff of different attributes when they shop online. This could be important as different attributes may mean different to various consumers, and investigating these differentiations between the valuation of attributes would allow retailers to specify their investments to maximize the returns on the different types of consumers. This investigation can be completed using the choice experiment section of our survey, allowing us to further understand the tradeoff between attributes of online grocery shopping overall and between different clusters of consumers. For instance, one of the most important attributes not investigated in the study is the pick-up and online delivery services offered. This difference affects consumer behavior and choice between shopping online and in-person and should be investigated as it could impact the clustering of individuals since our attitude questions directly refer to delivery while overlooking consumer attitudes towards being able to pick up their groceries.

Secondly, since we completed this investigation through an online Qualtrics survey, this is likely why the number of people who have used online grocery in the last six months is so

high. We have a biased sample since it is based on an online survey. Although collecting the data online seems to fit the theme of online grocery shopping, we naturally eliminated the potential offline users, who could be the potential market for online grocery in the future while simultaneously introducing the sample selection bias. This allows for results that provide insight to retailers on best practices for marketing online to consumers while lacking more information on marketing in traditional manners. A telephone survey to compare the results with the online survey would provide more information about the opinions and behaviors of online respondents. We do also have a higher amount of people who receive SNAP benefits than the average population as well, which indicates our sample may not reflect the general representative of an average grocery shopper in the US. These are the two major limitations of our investigation due to the time and funding restriction. However, the study provides insight that has been missing in the past and allows for a broader understanding of different clusters of consumers' perceptions towards online grocery shopping, such as SNAP recipients.

Conclusion

This study uses K-Means clustering to segregate a sample of 1,530 respondents into four clusters and employs logistic and ordinary least squares regression analysis to study the impact of consumer attitudes on their usage and expenditure within the online grocery market and in-person grocery shopping to advise retailers on demographic and household characteristics to target their marketing effectively. The results show that there is a group of consumers (around 15% of the sample) who love online grocery shopping and use it the most, with more than 90% of them having used online grocery shopping in the last six months and finding no barriers towards usage only indifference towards certain aspects. The tradeoff cluster (around 37% of the sample) has a high targetability by retailers as they see the temporal value of shopping for their

groceries online but are still hesitant due to the financial and trust-related aspects of online grocery shopping. There is a decent portion of the sample (around 22% of the sample) who is indifferent towards the three investigated aspects of online grocery shopping, who can be swayed easily as they only identify a single financial aspect as a negative to shopping for groceries online and by targeting this group with marketing that emphasizes the trustworthiness and temporal benefits to shopping online could be converted easily to users of OGS. Finally, there is a cluster of older folks (around 26% of the sample) who are haters of online grocery shopping as they see it as a costly and unnecessary method to obtain their groceries who would be difficult to convert into users and should be targeted last as more resources would need to be invested to convert as compared to the tradeoff and indifferent clusters of consumers. These results are primarily beneficial for retailers as they give insight into what targets of consumers should be targeted with marketing resources to demonstrate the temporal, financial, and trust-related benefits of shopping for groceries online using techniques targeted towards younger, educated, time-starved consumers who have experience with using online services already. The most consistent barrier amongst the groups is financial and trust-related issues with OGS, which can be addressed using marketing strategies to emphasize security to alleviate the concerns. Policymakers and retailers too, should note that more SNAP beneficiaries are adopting online grocery shopping thanks to the broader acceptance of SNAP benefits by online retailers. Therefore, finding a balanced way to subsidize online grocery shopping providers and encourage more SNAP recipients to utilize the shopping platforms may both support the reduction of our food desert and food security problem and meanwhile, provide more opportunities for it to become more affordable and accessible in the future through increased demand. Eventually, through demand-based improvements online grocery retailers can help improve social welfare

through increased healthy food purchases, create new jobs, and help those who need nutritious food access them conveniently and affordably.

List of References

Anesbury, Z., M. Nenycz-Thiel, J. Dawes, and R. Kennedy. “How do Shoppers Behavior Online? An Observational Study of Online Grocery Shopping.” *Journal of Consumer Behaviour* **15** (2016): 261-250.

Appelhans, B., E. Lynch, M. Martin, L. Nackers, V. Cail, and N. Woodrick. “Feasibility and Acceptability of Internet Grocery Service in an Urban Food Desert, Chicago, 2011-2012.” *Preventing Chronic Disease* **10** (2013).

Bauerova, K. “Consumers’ Decision-Making in Online Grocery Shopping: The Impact of Services Offered and Delivery Conditions.” *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* **66(6)** (2018): 1239-1247.

Benn, Y., T. Webb, B. Chang, and J. Reidy. “What Information Do Consumers Consider, and How Do They Look for it When Shopping for Groceries Online?” *Appetite*. **89** (2015): 265-273.

Bezirgani, A., U. Lachapelle. “Online Grocery Shopping for the Elderly in Quebec, Canada: The Role of Mobility Impediments and Past Online Shopping Experience.” *Travel Behaviour and Society* **25** (2021): 133-143.

Brand, C., T. Schwanen, J. Anable. “‘Online Omnivores’ or ‘Willing but Struggle’? Identifying Online Grocery Shopping Behavior Segments Using Attitude Theory.” *Journal of Retailing and Consumer Services* **57** (2020): 102195.

Breuglemans, E., and K. Campo. “Effectiveness of In-Store Displays in a Virtual Store Environment.” *Journal of Retailing* **87** (2011): 75-89.

Breuglemans, E., K. Campo, and E. Gijsbrechts. “Shelf Sequence and Proximity Effects on Online Grocery Choices.” *Market Letter* **18** (2007): 117-133.

Campo, K., E. Breuglemans. “Buying Groceries in Brick and Click Stores: Category Allocation Decision and the Moderating Effect of Online Buying Experience” *Journal of Interactive Marketing* **31**(2015):63-78.

de Magalhaes, D. “Analysis of Critical Factors Affecting the Final Decision-Making for Online Grocery Shopping.” *Research in Transportation Economics* **87**(2021): 101088.

Etumnu, C., and N. Widmar. “Grocery Shopping in the Digital Era.” *Choices* (2020):35(2)

Frank, D., and A. Peschel.. “Sweetening the Deal: The Ingredients that Drive Consumer Adoption of Online Grocery Shopping.” *Journal of Food Products Marketing* **26**(2020): 535-544.

Hand, C., F. Riley, P. Harris, J. Singh, and R. Rettie. “Online Grocery Shopping: The Influence of Situational Factors.” *European Journal of Marketing* **43**(2008): 1206-1219.

Hanus, Gabriela. “CONSUMER BEHAVIOUR DURING ONLINE GROCERY SHOPPING.” *CBU* **4**(2016):010-013.

Harris, P., F. Riley, D. Riley, and C. Hand. “Online and Store Patronage: A Typology of Grocery Shoppers.” *International Journal of Retail Distribution and Management* 45(2017): 419-446.

Harris-Lagoudakis, K. “Online shopping and the healthfulness of grocery purchases.” *American Journal of Agricultural Economics* 104(2022): 1050-1076.

Huyghe, E., J. Verstraeten, M. Geuens, and A. van Kerckhove. “Clicks as a Healthy Alternative to Bricks: How Online Grocery Shopping Reduces Vice Purchases.” *Journal of Marketing Research* 54(2017): 61-74.

IBISWorld—Industry Statistics - United States. Internet site:
<https://www.ibisworld.com/industry-statistics/market-size/online-grocery-sales- united-states/>
(Accessed May 5, 2023).

IBISWorld—Industry Statistics - United States. Internet site:
<https://www.ibisworld.com/industry-statistics/number-of-businesses/online- grocery-sales- united-states/> (Accessed May 5, 2023).

Kühn, F., Lichters, M., & Krey, N. The touchy issue of produce: Need for touch in online grocery retailing. *Journal of Business Research*. **117**(2020): 244-255.
<https://doi.org/10.1016/j.jbusres.2020.05.017>.

Lennon, S., Y. Ha, K. Johnson, C. Jasper, M. Damhorst, and N. Lyons. “Rural Consumers’ Online Shopping for Food and Fiber Products as a Form of Outshopping.” *Clothing and Textiles Research Journal* 27(2009): 3-30.

Martinez, O., B. Taliaferro, N. Rodriguez, J. Athens, C. Abrams, and B. Elbel. “EBT Payment for Online Grocery Orders: A Mixed-Methods Study to Understand its Uptake Among SNAP Recipients and the Barriers to and Motivators for its Use.” *Journal of Nutrition Education and Behavior* 50(4): 396-402 (2018).

Melis, K., K. Campo, E. Breugelmans, and L. Lamey. “The Impact of the Multi-channel Retail Mix on Online Store Choice: Does Online Experience Matter?” *Journal of Retailing* 91 (2015): 272-288.

Mortimer, G., S. e Hasan, L. Andrews, and J. Martin. “Online Grocery Shopping: The Impact of Shopping Frequency on Perceived Risk.” *The International Review of Retail, Distribution, and Consumer Research* 26(2016): 202-2243.

Pitts, S., S. Ng, J. Blitstein, A. Gustafson, and M. Niculescu. “Online Grocery Shopping: Promise and Pitfalls for Healthier Food and Beverage Purchases.” *Public Health Nutrition* 21(2018): 3360-3376.

Robinson, H., F. Riley, G. Rolls-Willson. “The Role of Situational Variables in Online Grocery Shopping in the UK.” *The Marketing Review* 78:9-107 (2007).

Rogus, S., J. Guthrie, M. Niculescu, L. Mancino. “Online Grocery Shopping Knowledge, Attitudes, and Behaviors Among SNAP Participants.” *Journal of Nutrition Education and Behavior* 52(5): 539-545 (2019).

Saphores. “E-Shopping Changes and the State of E-Grocery Shopping in the US - Evidence from National Travel and Time Use Surveys.” *Research and Transportation Economics* 87(2021).

Singh, R., M. Soderlund. “Extending the Experience Construct: An Examination of Online Grocery Shopping.” *European Journal of Marketing* 54(2019): 2419- 2446.

Statista—Online Grocery Shopping - United States. Internet site:
<https://www.statista.com/topics/1915/online-grocery-shopping-in-the-united-states/#topicOverview> (Accessed May 5, 2023).

van Ewijk, B., J. Steenkamp, and E. Gijsbrechts. “The Rise of Online Grocery Shopping in China: Which Brands Will Benefit?” *Journal of International Marketing* 28(2) (2020): 20-39.

Yenerall, J., You, W., & Hill, J. (2020). Beyond the supermarket: Analyzing household shopping trip patterns that include food at home and away from home retailers. *BMC Public Health*, 20(1), 1747. <https://doi.org/10.1186/s12889-020-09882-0>

Yongchang Wei et al. "Online Purchase Intention of Fruits: Antecedents in an Integrated Model Based on Technology Acceptance Model and Perceived Risk Theory." *Frontiers in Psychology*, 9 (2018). <https://doi.org/10.3389/fpsyg.2018.01521>.

Zatz, Laura Y et al. “Comparing Online and In-Store Grocery Purchases.” *Journal of Nutrition, Education, and Behavior* 53.6 (2021): 471-479.

Vita

Harrison Clark's educational journey began at the University of Tennessee-Knoxville, where he embarked on a Master's program in Agricultural Economics from August 2022 to May 2024. Before this, he completed his Bachelor of Arts in Economics at the same institution, graduating Magna Cum Laude in May 2022. Clark's academic foundation was further solidified with an Associate in Science degree from Trident Technical College, achieved in July 2020, and a High School Diploma from Palmetto Scholars Academy, also obtained in July 2020, both with Magna Cum Laude honors.

Clark's primary research interests lie in the dynamics of Online Farmer's Markets, Online Grocery Stores, Online Sales, and Consumer Behavior. He also holds a secondary interest in Nutrition and Food Security, Health and Crime Economics, and Food Marketing.

Clark's scholarly contributions include a mix of peer-reviewed publications and conference papers. Notably, he co-edited a publication on the "Economic Impact of Tennessee Forest Product Exports in 2022" and submitted an article on the optimization of edible gardens for backyard producers. His conference presentations span topics from the market value of eco-labeled crops to the profitability of online farm product sales, reflecting his diverse interests within the field of agricultural economics.

Clark served as a Research Assistant for two months between June and August 2022, under the guidance of Professor Andrew Muhammad. This role involved investigating the agricultural trade relationship between the US and Brazil, leading to significant findings presented to the Economic Research Service. Additionally, he engaged in Applied Economic Analysis for three months under Senior Lecturer Benjamin Compton, tackling projects ranging from macroeconomic analysis to econometric analysis of public policy initiatives.

Clark has demonstrated proficiency in a wide range of technical and interpersonal skills. He is skilled in Microsoft products, Python for data analysis, and has a basic operation knowledge of Stata and R software. His academic diligence is evident from his consistent Magna Cum Laude graduations and early completion of an Associate degree. Clark's communication skills, leadership, and learning capabilities have been honed through various experiences, including achieving the Eagle Scout rank in the Boy Scouts of America. Harrison Clark's curriculum vitae showcases a committed and multifaceted professional with a strong foundation in agricultural economics, poised to contribute significantly to his field of interest.