INFLUENTIAL FACTORS ON THE LIKELIHOOD TO PURCHASE ALCOHOL ONLINE

Amanda G. Schroeder

An honors thesis submitted to the faculty of the Kenan-Flagler Business School at the University of North Carolina at Chapel Hill

Chapel Hill
2018

Approved by __________________________
(Dr. Wendell Gilland)
ABSTRACT
Amanda G. Schroeder
Influential Factors on The Likelihood to Purchase Alcohol Online
(Under the direction of Dr. Wendell Gilland)

A 2016 study found that e-commerce businesses in the United States generated $322.17 billion in sales (Johnson, 2017), and other estimates predict that the e-commerce industry will grow by 23.2% in 2017 (Chaffey, 2017). While research concurs that e-commerce is a growing industry, other studies aimed to understand consumer behavior online have presented conflicting conclusions. My research seeks to identify factors that influence consumers’ buying intent in an emerging industry: e-commerce for buying alcohol.

To identify influential factors on the likelihood to buy alcohol online, a consumer behavior survey was conducted in the United States. The survey asked respondents to provide demographic information, as well as information on their current e-commerce behavior and alcohol consumption habits. Cross tabulations and logistic regression analyses were performed to identify influential factors on a respondents’ likelihood to use e-commerce for buying alcohol. Findings indicate that the following factors significantly influence the intent to buy alcohol online: (1) the frequency a respondent currently buys items online, (2) how often a respondent buys alcohol, (3) how long a respondent expects regular online orders to arrive, (4) how many days after purchase a respondent waits to consume an alcoholic purchase, and (5) the type of alcoholic beverage respondents consume the most.
ACKNOWLEDGEMENTS

The completion of this thesis would not have been possible without the unconditional support of the following individuals. I’d like to recognize their contributions because without them, this would not be possible.

Dr. Wendell Gilland – thank you for being an incredibly reliable mentor throughout this process. Without your advice and guidance, this thesis would not be where it is today.

Dr. Nick Didow & Dr. Mahka Moeen – thank you for serving on my committee, for engaging in my work, and for providing thoughtful feedback.

Dr. Patty Harms – thank you for instilling the confidence in me to complete this thesis. Without your help, I would not have been able to build a strong enough foundation to complete my thesis.

My parents – thank you for always being proud of me. Your love and support are what gives me the strength to lean into any challenge. Because of you, I’ve learned to work hard and stay humble.

Colin – thank you for always empowering me in the moments where I feel most inadequate.

This project was supported by the Tom and Elizabeth Long Excellence Fund for Honors administered by Honors Carolina.
# TABLE OF CONTENTS

ABSTRACT .............................................................................................................. ii

ACKNOWLEDGEMENTS ....................................................................................... iii

LIST OF TABLES ...................................................................................................... vi

LIST OF FIGURES ................................................................................................... vii

Chapter

I. INTRODUCTION ................................................................................................. 1

II. LITERATURE REVIEW ...................................................................................... 3

   A. Channel Strategy ............................................................................................. 3
   B. E-Commerce ..................................................................................................... 8
   C. The Alcohol Industry ...................................................................................... 11
   D. Gap in the Narrative ...................................................................................... 16

III. METHODOLOGY ............................................................................................... 17

   A. Research Design ............................................................................................. 17
   B. Gathering Data ............................................................................................... 18
   C. Key Metrics ..................................................................................................... 19
   D. Limitations ...................................................................................................... 22
IV. RESULTS ................................................................. 23
   A. Demographics .................................................. 24
   B. Cross Tabulations ............................................. 25
   C. Logistic Regression Analyses ............................... 34
V. IMPLICATIONS ....................................................... 42
   A. Implications of Cross Tabulations ......................... 42
   B. Implications of Logistic Regression Results ............. 45
VI. CONCLUSION .......................................................... 48
VII. APPENDIX
   A. Survey Consent Form ........................................... 50
   B. Survey .............................................................. 51
VIII. REFERENCES ........................................................ 57
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>Logistic regression model variables defined</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Age distribution of sample population versus United States population</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Distribution of sample population’s education attainment level</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Distribution of likelihood indicators by income level</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Sample population’s shipping expectations for alcohol orders online</td>
</tr>
<tr>
<td>Table 4.5</td>
<td>Logistic regression results from original model including all 13 independent variables</td>
</tr>
<tr>
<td>Table 4.6</td>
<td>List of logistic regression iterations by variable removed</td>
</tr>
<tr>
<td>Table 4.7</td>
<td>Logistic regression results from final reduced model</td>
</tr>
<tr>
<td>Table 4.8</td>
<td>Classification table from final reduced logistic regression</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The channel continuum</td>
<td>5</td>
</tr>
<tr>
<td>2.2</td>
<td>The three-tiered distribution system diagram</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>The alcohol distribution channel with e-commerce</td>
<td>14</td>
</tr>
<tr>
<td>4.1</td>
<td>Distribution of likelihood indicators in each age bracket</td>
<td>26</td>
</tr>
<tr>
<td>4.2</td>
<td>Distribution of likelihood indicators by gender</td>
<td>27</td>
</tr>
<tr>
<td>4.3</td>
<td>Distribution of likelihood indicators by education attainment level</td>
<td>28</td>
</tr>
<tr>
<td>4.4</td>
<td>Distribution of likelihood indicators by employment status</td>
<td>29</td>
</tr>
<tr>
<td>4.5</td>
<td>Distribution of likelihood indicators by e-commerce shipping preferences</td>
<td>31</td>
</tr>
<tr>
<td>4.6</td>
<td>Distribution of likelihood indicators by reason respondent uses e-commerce</td>
<td>32</td>
</tr>
<tr>
<td>4.7</td>
<td>Distribution of the type of alcoholic beverage the sample population consumes most versus the United States Population</td>
<td>33</td>
</tr>
</tbody>
</table>
INTRODUCTION

Since the dot-com era in the late 1990’s – early 2000’s, the internet has revolutionized the way people find information, stay connected with others, and live our day-to-day lives. The internet has also transformed the way people buy and sell goods and services. Electronic commerce (e-commerce) has created new businesses and industries, while destroying others. Since 2014, e-commerce share of global GDP has tripled from 0.54% to 1.61%, and e-commerce sales in the United States reached $580 billion in 2017 (Moagar-Poladian, Dumitrescu, & Tanase, 2017, p. 174).

As the e-commerce industry continues to grow, researchers have aimed to discover patterns of how consumers behave online. Unfortunately, research has failed to establish a consensus on what factors influence e-commerce buying behaviors. Thus, conducting research on consumer behavior at an industry level can help companies develop targeted sales and marketing strategies.

Despite unclear answers on how consumers behave online, many businesses have found success using e-commerce. Amazon, founded in 1995, started as an online bookstore and has since diversified to selling thousands of products, becoming the largest e-commerce platform by sales revenue in the world (Moagar-Poladian et al., 2017, p. 171). More recently, Amazon acquired Whole Foods, thus signaling a future where grocery delivery is the norm (Turner, Wang, & Soper, 2017). Yet there still exists one industry that has yet to widely adopt e-commerce: the alcohol industry. Strict distribution
laws in the United States, enacted post-Prohibition, have made adopting e-commerce more complicated.

The goal of this research is to identify factors that influence the likelihood to use e-commerce to buy alcohol. Factors included in this research are demographic variables, as well as, their current online buying and alcohol consumption behavior. Through a consumer behavior survey of adults across the United States, I have attempted to draw meaningful conclusions from regression analyses to identify potential influencing factors. I’m hopeful that the correlations found may allow companies in the alcohol industry to make more meaningful and targeted approaches to developing their e-commerce strategies.

13 variables were selected to identify influential factors on the likelihood of a respondent to buy alcohol online. Only five variables indicated statistical significance: (1) the frequency a respondent currently buys items online, (2) how often a respondent buys alcohol, (3) how long a respondent expects regular online orders to arrive, (4) how many days after purchase a respondent waits to consume an alcoholic purchase, and (5) the type of alcoholic beverage respondents consume the most.
LITERATURE REVIEW

The benefits of e-commerce have become a predominant source of competitive advantage in a variety of industries and has particularly disrupted channel strategy. In the following section I will review relevant literature that provide a useful background for my research. I will first explain channel strategy, and how e-commerce has evolved throughout the past two decades. Then I will discuss relevant research on the alcohol industry and insights from an industry professional.

Channel Strategy

Channel is broadly defined as the means a company goes to market with a product or service (Friedman & Furey, 1999, p. 3). Historically, business managers have prioritized improving brand recognition, product innovation, and efficient supply chains in order to achieve competitive advantage. However, as companies and industries face more saturated markets, optimizing channel strategy has become a priority for many.

While brand improvement, product innovation, and lean supply chains are vital factors for success, channel strategists have recognized that “in areas where there are few product differences… the key differentiating value often comes from the means of acquisition – the channel” (DeVincentis & Rackham, 1998, p. 222). Particularly with products that are mass-produced and in highly-saturated markets, channel innovation is beginning to separate the winners and losers in a market.
For example, Amazon has proven that channel innovation has the power to completely disrupt the fabric of an industry. For decades, consumers went to physical retailers to purchase a book. Yet in 1995, Amazon adopted a unique channel in the industry – e-commerce. By developing an e-commerce channel strategy, Amazon has been able to grow exponentially by engineering faster checkout online, offering convenient shipping methods, and data analytics to track consumer behavior (Perreault, Cannon & McCarthy, 2017, p. 57-8). Not only did Amazon help introduce e-commerce to the bookstore industry, they also partnered and/or acquired e-commerce businesses in other sectors (i.e. pharmacies, pet supplies, home-goods, and auctions) to expand their offerings. As the company began to expand beyond the bookstore industry, analysts coined the phrase “Amazon-ing a sector.” This phrase originated after Amazon became so dominant in the bookstore industry, that it became difficult for others to succeed. Researchers believe that Amazon’s innovative use of e-commerce, in part, led to the bankruptcy of Borders bookstore in 2011 (Chaffey, 2015, p. 645). Amazon is proof that a focus on channel innovation has the capacity to completely disrupt multiple industries and can transform the way consumers find, purchase, and acquire products.

Different types of channels. Channels tend to fall into one of three categories: direct marketing, indirect marketing, and direct sales. Each category has varying levels of value-added to the sale and cost per transaction. Figure 2.1 illustrates each type of channel.
Figure 2.1. The channel continuum. Source: The Channel Advantage (Friedman & Furey, 1999, p. 46)

Direct marketing channels have historically featured low levels of value-added to the sale and low costs per transaction. However, more recent research has shown that e-commerce companies are able to add value to the consumer through features like fast delivery, competitive prices, and more product options (Chaffey, 2015). Companies and industries that choose direct marketing channels sell direct to consumer and generally have low touch relationships with their consumer. Historically, low touch relationships have been defined by offering little to no interaction between the seller and buyer (Friedman & Furey, 1999, p. 47). However, the internet category deviates from this definition. In fact, the internet offers both consumers and businesses more touchpoints. Meaning, the internet allows consumers to discover more businesses online and gives businesses the ability to expand their target market reach (Chaffey, 2015, p. 42). Other channels in this category include telemarketing businesses and retail stores that sell direct to consumer and require little customization for consumers. For example, a direct
Marketing channel for Nike would mean selling their products at a Nike retail store, as opposed to selling in a department store.

Indirect marketing channels are also referred to as the reseller structure, because companies sell their product through an intermediary (Friedman & Furey, 1999, p. 47). One popular example of an industry that uses the indirect marketing channel is the alcohol industry. Manufacturers, like Anheuser-Busch InBev, sell to distributors/wholesalers. Those distributors then sell to accounts, like restaurants and grocery stores. This example is unique because it features two intermediaries between the manufacturer and end consumer. Specific channels in this category include distributors, value-added partners, and other retail stores. An example of an indirect marketing channel in another industry is Mars, Incorporated. An indirect channel for Mars, Inc. would mean selling their iconic M&M’s at grocery stores, instead of direct to consumers at an M&M’s World retail store.

Direct sales channels use field sales forces and feature high cost per transaction and high levels of added-value for the consumer (Friedman & Furey, 1999, p. 48). Companies and industries that use this channel generally conduct business to business sales that involve complex products or services. Direct sales channels also have high touch relationships with their customer. High touch relationships require high levels of interaction between buyer and seller (Friedman & Furey, 1999, p. 47).

How companies pick the right channel. With a variety of different channel options, companies must understand how to pick the best channel for their product and their customer. The Channel Advantage (1999) provides a three-pronged framework for
determining an optimal channel strategy: (1) define product-market focus, (2) understand how customers want to buy, and (3) determine the complexity of a product.

The goal of product-market focus is to segment a business’ revenues to understand the following: which customers should a company target to maximize revenues and what markets are those customers in (Friedman & Furey, 1999, p. 18). For example, Canon sells a variety of products from cameras and lenses to printers and copiers. So, the product-market focus for Canon means targeting lens and camera products to individual users and targeting printers and copiers to large corporate accounts. (Friedman & Furey, 1999, p. 18).

A company should also uncover how a customer wants to buy. Businesses can understand how their customers want to buy through determining their personal preferences, behaviors, and buying criteria (Friedman & Furey, 1999, p. 30). For example, do customers prioritize price or quality of a product? Where and when do customers purchase a product? Analytics on consumer behavior can be helpful to understand what factors influence buying behaviors (Perreault, Cannon & McCarthy, 2017, p. 57-8).

Companies must also understand the complexity of their product or service. For example, if a customer requires consultation and support, the product or service is complex. Complex products and services require high-touch/direct sales channels. If the product is not complex, companies can utilize low-touch/direct marketing channels (Friedman & Furey, 1999, p. 55).

Friedman and Furey (1999) conclude The Channel Advantage by arguing that companies may need a mix of channels to optimize their strategy (p. 174). This trend is
widely used in the fashion industry, where apparel brands utilize their own retail stores (direct marketing channel) and department stores (indirect marketing channel) to sell products.

E-Commerce

_E-commerce_ is a platform that defines “all electronically mediated transactions between an organization and any third party,” where people can buy and sell products and services online (Chaffey, 2015, p. 13). Within the direct marketing channel category, channel strategists have begun to notice the increasing use of the internet. In 2016, e-commerce businesses in the United States generated $322.17 billion in sales (Johnson, 2017). In 2017, analysts predict that sales from the e-commerce sector will grow 23.2% (Chaffey, 2017). While the internet was initially considered an advertising tool, e-commerce is now an economically viable way for businesses to increase their consumer reach, reduce supply chain costs, and adapt more quickly to market demands (Chaffey, 2015, p. 31).

**Who sells through e-commerce.** Companies that use e-commerce generally sell non-complex, off-the-shelf products that are mass-produced and ready for use, without the need of customization or support from the seller. Companies that fall within this category experience both benefits and costs when using e-commerce.

E-commerce can be beneficial financially and help broaden potential customer bases. E-commerce also offers lower cost per transaction, mainly because companies no longer incur inventory and labor costs at brick and mortar stores. Online stores are also
accessible to more people. Therefore, higher volume of transactions can help reduce the cost per transaction (Friedman & Furey, 1999, p. 153).

E-commerce is also beneficial beyond financial metrics. Online platforms often allow companies to collect and analyze market research, to better inform business decisions. For example, tracking an individual’s buying patterns can help companies create a targeted selling approach for individual customers (Friedman & Furey, 1999, p. 154).

Despite the benefits, e-commerce also has disadvantages. The largest barrier to entry for implementing e-commerce is the financial investment. Developing a website, readjusting procurement, and supply chain operations, is costly (Chaffey, 2015). Capital is also needed to advertise and drive consumers to buy online. Beyond the financial burden, companies with e-commerce platforms struggle with new competition. As e-commerce makes it easier for consumers to find more options online, businesses have to compete with more companies (Chaffey, 2015, p. 42). Another significant disadvantage with e-commerce is understanding what their customers want on an e-commerce website. In 2009, researcher Dave Chaffey completed a case study to understand why businesses failed during the Dot-Com era. Chaffey concluded that while the failed businesses were largely founded on innovative ideas, they required a shift of consumer behavior (Chaffey, 2015, p. 70). So, while these failed companies focused on developing technology infrastructure, most failed to convince their buyers to buy online.

**Who buys through e-commerce.** As e-commerce continues to grow in countries across the globe, researchers have yet to clearly identify who buys online, what people value on e-commerce platforms, and how those values influence online buying behavior.
A study on the evolution of retail e-commerce in 2016 concluded that 82.5% of Americans have access to the internet and 77% of Americans are in the digital buyer category, meaning they have and intend to continue to purchase items online. American consumers in the digital buyer category spent over $157 billion on e-commerce websites in 2016 (Moagar-Poladian et al., 2017, p. 176-7).

While research agrees that a majority of Americans shop online, researchers are unclear which demographic characteristics have an effect on online shopping behavior. In a study to uncover influential factors for commodity purchases online, Dillon and Reif (2004) concluded that younger people are more comfortable with technology, and therefore are more likely to purchase items online (p. 12). However, another study by Hui and Wan (2009) concluded that age did not have a relationship with the likelihood of buying groceries online (p. 1481).

Beyond studying demographic factors, researchers have attempted to understand what e-commerce shoppers value. Prashar, Vijay, and Parsad (2017) concluded that consumers are either *hedonic shoppers*, who value web design and ease of use, or *utilitarian shoppers*, who value price and quality (p. 12). Depending on what customers value, companies can tailor their e-commerce selling approach. Hjort, Lantz, Ericsson, and Gattorna (2013) suggest that companies will have customers that value different characteristics. Therefore, companies need to segment their shoppers and individually target customers depending on what they value.

Other researchers have attempted to uncover which specific characteristics of e-commerce are important to buyers. One study on online textbook sales at James Madison University concluded that quality and price were the most important factors to
consumers, while convenience and availability weren’t as valued (Dillon & Reif, 2004, p. 11-2). A more recent study in 2011 suggested that the convenience of saving time and finding the best products were the most influential factors when shopping online (Punj, 2011, pp. 140). Cleary, research has not been able to unanimously make conclusions about online buying preferences and behaviors.

The Alcohol Industry

**Background and key terms.** The history of the alcohol industry plays a large role in understanding how the industry is structured today. Early temperance groups in the 1850’s held that a prohibition of alcohol would resolve an assortment of social issues in the United States: drunken workers/servants, unemployment, poverty, and crime (Levine & Reinarman, 1991, p. 464). After gaining more popularity throughout the late 1800’s, prohibitionist groups united under the Anti-Saloon League. In the early decades of the 20th century, the League hired lawyers to write prohibition laws, eventually lobbying legislators in Congress for a constitutional prohibition of alcohol. By January of 1920, the 18th amendment of the United States constitution took effect, banning the manufacture, sale, transportation, import, and export of alcohol (Levine & Reinarman, 1991, p. 463). Despite the constitutional amendment, alcohol in the United States was still widely available across the country during the prohibition era. Illegal manufacturers and sellers, called bootleggers and speak-easies, began selling a variety of alcoholic beverages, often manufactured with higher levels of alcohol by volume (ABV) (Levine & Reinarman, 1991, p. 464).
By the 1930’s, militant groups and mobs threatened revolts in an effort to repeal prohibition laws. In November of 1932, the United States senate voted to enact the 21st amendment to the United States constitution, repealing the 18th amendment. Following the prohibition era, federal and state governments recognized the need to regulate and tax the alcohol industry to ensure bootleggers and speakeasies seize the production of unsafe and illegal alcoholic beverages (Levine & Reinarman, 1991, p. 466). John D. Rockefeller and Raymond Fosdick published Toward Liquor Control in various newspapers and magazines, which laid out central laws and regulations that are still intact in the US today. Alcohol is regulated at both the state and federal-levels. Examples of state-level regulations include state owned and run liquor stores. For states that haven’t established governmentally owned liquor stores, the federal government regulates privately owned stores through licensing.

Government regulation is also present throughout the alcohol distribution channel (Levine & Reinarman, 1991, p. 476). Coined the three-tiered system, manufacturers sell products to distributors/wholesalers. Product is then sold into retail accounts, like grocery stores, restaurants, and liquor stores, where the end consumer can purchase products.

The three-tiered distribution system falls under the indirect marketing channel category, which is illustrated in Figure 2.2.

![Figure 2.2. The three-tiered distribution system diagram. Adapted from: Essentials of Marketing (Perreault, Cannon & McCarthy, 2017, p. 307).](image-url)
**Industry specific research.** Roland Zullo, professor of privatization of business, tested the efficiency of alcohol laws and regulations. Zullo’s (2017) research studied the effectiveness of governmentally owned liquor stores. Zullo used the New Public Management theory, which argues that privatization of markets maximizes public value (Zullo, 2017, p. 190). Public value was defined by price, availability of stores, product assortment, and store hours. Ultimately, Zullo found that a mixture of private and governmentally owned stores maximizes public value (Zullo, 2017, p. 204). For example, benefits of privatized stores include easier access to liquors/spirits, more store locations, and longer store hours. However, governmentally owned stores provided consumers with larger store sizes and wider assortments of products (Zullo, 2017, p. 191). So, Zullo concluded that a mix of private and governmentally owned stores is optimal because individuals value the benefits of each store differently. However, Zullo is unclear what exactly the mix might look like. Zullo’s results align with the notion from *The Channel Advantage* that companies and industries may need a mix of channels to maximize company and consumer value (Friedman & Furey, 1999, p. 174).

**Insights from industry professional.** To get a more robust understanding on the current state of e-commerce in the alcohol industry, I gathered information from the Global E-commerce Director of North America at Anheuser-Busch InBev, Carolyn Littlefield (C. Littlefield, phone interview, Feb. 9, 2018). Throughout our conversation, I attempted to understand the following:

- How will e-commerce change the three-tiered distribution channel?
- How are companies in the alcohol industry adjusting towards adopting e-commerce?
How are alcohol e-commerce companies using data to inform marketing and sales decisions?

*How will e-commerce change the three-tiered distribution channel?* Ms. Littlefield first explained where e-commerce fits into the three-tiered distribution system. The standard alcohol distribution channel is illustrated in Figure 2.2, and begins with a producer, like Anheuser-Busch InBev, who’s product is sold through wholesalers and retailers before reaching the end consumer. Ms. Littlefield explained that e-commerce now serves as an intermediary between the retailer and consumer (Illustrated in Figure 2.3).

![Diagram of the alcohol distribution channel with e-commerce](image)

*Figure 2.3. The alcohol distribution channel with e-commerce*

So, instead of a consumer purchasing alcohol in a retail location, like a grocery, liquor, or convenience store, the customer would place an order through an e-commerce platform, that would fulfill the order at a retail store and deliver the order to the consumer.

*How are companies in the alcohol industry adjusting towards adopting e-commerce?* One adjustment towards e-commerce has involved alcohol companies educating wholesalers on the potential impact of an e-commerce platform. While producers and wholesalers are generally separate companies (some wholesalers are wholly-owned by producers), Ms. Littlefield explained that Anheuser-Busch InBev has been working to educate wholesalers on how e-commerce effects the three-tiered distribution system. E-commerce platforms that sell alcohol often provide a larger
assortment for consumers, so wholesalers may experience temporary growing pains as consumers demand larger assortments of products. Another adjustment for alcohol companies (manufacturers, wholesalers, and retailers) has meant forming relationships with existing e-commerce platforms (i.e. Drizly, Minibar, and Amazon PrimeNow). Ms. Littlefield explained that although there is a lot of grey area with the nature of relationships between e-commerce platforms and producers, Anheuser-Busch InBev is excited to continue leveraging e-commerce platforms to market their products.

**How are alcohol e-commerce companies using data to inform marketing and sales decisions?** As previously stated, data collection on consumers is an advantage for e-commerce businesses. The collection of consumer data, like search engine keywords, purchase transactions, and browsing history allows businesses to create more targeted marketing approaches (Chaffey, 2015, p. 535). However, because selling alcohol through e-commerce platforms is a relatively new industry, data has gone uncollected and unanalyzed. For example, Anheuser-Busch InBev works closely with a market research company called the Industrial Research Institute (IRI). Ms. Littlefield explained that as of 2018, the IRI has yet to establish sophisticated metrics to track e-commerce activity in the alcohol industry. For example, sales at the retail level can be to an e-commerce platform or to a consumer, but the IRI has no way of distinguishing the two. Ms. Littlefield also explained that while Anheuser-Busch InBev has conducted preliminary research about consumer behavior with buying alcohol online, there are still a lot of open questions.
**Gap in Narrative**

Research in this literature review presents a gap that my research attempts to supplement. While recent research agrees that e-commerce will continue to grow in the coming years, other studies present conflicting conclusions as to which factors influence buying behavior online. As the data suggests that consumer preferences vary from person to person and across industries, the gap in the narrative presents an opportunity to conduct e-commerce consumer behavior research on an industry-specific level. Further, since platforms that sell alcohol online are still an emerging trend in the e-commerce industry, there is little to no research on how consumers behave buying alcohol online. So, my research on identifying influential factors on the likelihood to use e-commerce to buy can help elevate the industry.
METHODOLOGY

To determine influential factors on the likelihood to buy alcohol online, I have conducted market research through a survey of adults in the United States. This research aimed to assess respondent’s current online buying and alcohol consumption habits to identify which variables influence the likelihood to buy alcohol online. In the following section I will describe my research design, methods for gathering data, key metrics used to interpret the data, and limitations to my approach.

Research Design

The survey began with a consent form (Appendix A) that described the purpose of the survey, the estimated completion time, and stated the respondent’s rights to discontinue the survey at any time. Upon agreeing to the consent form, respondents were asked to disclose basic demographic information: age, gender, education attainment level, employment status, and income level. Obtaining this demographic information was crucial to ascertain if certain demographics are prone to specific buying habits.

After requesting demographic information, the survey prompted a series of questions to gather information about respondents’ current online shopping behavior: (1) how often do you buy items online, (2) why do you buy online versus in-store, (3) how quickly do you expect your order to arrive, and (4) what are your preferences on paid or free shipping. See Appendix B for survey questions and answer options.
Following questions about online buying behavior, respondents were asked a series of questions regarding their alcohol consumption habits. Questions in this section of the survey aimed to gather information about (1) how often respondents drink alcohol, (2) what beverages they prefer, (3) where they are likely to purchase alcohol, and (4) how quickly they consume alcohol products after purchase.

The final section of my survey was specifically targeted to understand the respondent’s thoughts on e-commerce platforms that sold alcohol. Questions initially asked (1) whether they had heard of buying alcohol online, (2) indicate if they are likely to use alcohol e-commerce platforms, and (3) what type of beverage they are most likely to purchase online. The survey then prompted the following:

*As of February 2018, alcohol deliveries require that an adult 21 years or older be present at the time to sign for the delivery and cannot be left at a residence without a signature.*

Following this prompt, respondents were asked again how likely they are to use e-commerce platforms to buy alcohol and were prompted to check all limitations or concerns that may prohibit their likelihood to use e-commerce to buy alcohol. Respondents were also provided a free-response option to write-in other limitations not mentioned in the survey.

**Gathering Data**

My survey was certified by the Institutional Review Board (IRB) at The University of North Carolina at Chapel Hill on February 13th, 2018. My market research survey was designed and distributed through Survey Monkey, an online survey tool. This platform was chosen because it offers a tool called Audience, which allowed me to easily
distribute the survey to a large number of people and enforce a minimum age of respondents. Per the IRB’s ethical guidelines, my market research could only be sent to people at or above the legal drinking age (21 years) in the United States. If sent to people younger than 21 years old, respondents may have disclosed potentially illegal information about their alcohol consumption habits. So, Survey Monkey’s Audience tool allowed me to control that the survey be sent only to those 21 years or older. This online survey tool also allowed me to distribute my market research survey to people across the United States, ranging in age, education, and income levels. My ability to collect responses from a wide range of demographic variables was crucial in allowing me to draw meaningful inferences and trends.

**Key Metrics**

To draw significant conclusions from my survey, I began my analysis by identifying key statistical measures to assess the validity of my sample group. I did this by comparing the sample’s distribution of demographic variables (i.e. age, gender, education attainment level, employment status, and annual income level) with United States census information.

Before conducting sophisticated logistic regression analyses, I calculated various cross tabulations to examine preliminary trends in the distribution of likelihood by responses to various questions.

To understand the marginal effect of influential variables on the likelihood to buy alcohol online, I decided to use binary logistic regression models. This type of regression analysis was chosen because it allows researchers to create predictive models
using multiple independent variables on a single binary outcome (Robinson, n.d.). While many variables could have been used in the analysis, I decided to use 13 independent variables, which are defined in Table 3.1. The binary dependent variable for my analysis is Likelihood, which is also defined in Table 3.1. These variables were chosen based on research detailed in literature review. I began by conducting a binary logistic regression model including all 13 independent variables on the dependent variable, Likelihood. I then used a method of model reduction, where I removed statistically insignificant independent variables with the highest $p$-value, one at a time, until only statistically significant independent variables remained. A new binary logistic regression model was calculated 8 times before producing the final reduced model where all independent variables were significant.

Table 3.1

*Logistic regression model variables defined*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Response to Question 1: Specific age of respondent</td>
</tr>
<tr>
<td>Female</td>
<td>Response to Question 2: If respondent is female</td>
</tr>
<tr>
<td>Bach/Grad/PhD</td>
<td>Response to Question 3: If respondent’s highest education attainment level is either Bachelor’s, Graduate, or PhD</td>
</tr>
<tr>
<td>Part/Full-Time Employed</td>
<td>Response to Question 4: If respondent is either employed part-time or full-time</td>
</tr>
</tbody>
</table>
### Table 3.1

**Logistic regression model variables defined (continued)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>Response to Question 5: Annual income of respondent, where the midpoint of each income category was used</td>
</tr>
<tr>
<td>Frequent Shoppers</td>
<td>Response to Question 6: If respondent buys items online more than 4 times a month</td>
</tr>
<tr>
<td>Longer Delivery Expectation</td>
<td>Response to Question 8: If respondent doesn’t expect online orders to arrive until after 3 days</td>
</tr>
<tr>
<td>Frequent Drinkers</td>
<td>Response to Question 10: If respondent consumes alcohol more than 4 times a week</td>
</tr>
<tr>
<td>Wine</td>
<td>Response to Question 11: If respondent usually drinks wine more than liquor/spirits/mixed drinks and beer</td>
</tr>
<tr>
<td>Liquor</td>
<td>Response to Question 11: If respondent usually drinks liquor/spirits/mixed drinks more than wine and beer</td>
</tr>
<tr>
<td>Store</td>
<td>Response to Question 12: If respondent usually purchases alcohol at grocery, liquor, or convenience stores</td>
</tr>
<tr>
<td>Consume Immediately</td>
<td>Response to Question 13: If respondent usually consumes alcoholic purchases within 2 days</td>
</tr>
<tr>
<td>Frequent Alcohol Buyers</td>
<td>Response to Question 14: If respondent usually buys alcohol more than once a week</td>
</tr>
<tr>
<td>Likelihood</td>
<td>Response to Question 17: If a respondent indicated they are likely or somewhat likely to use e-commerce to buy alcohol</td>
</tr>
</tbody>
</table>
Limitations

There are several limitations with my research. The consumer behavior survey specifically presented time and resource limitations. For example, I was only able to distribute my survey to 369 people and had to omit 97 responses from people who don’t drink alcohol. While demographics within the sample group were relatively representative of the United States population, having a larger sample size may alter the results. Another limitation is that the survey was only distributed to consumers in the United States. So, the findings in my research do not reflect global trends of likelihood to use e-commerce to buy alcohol.

The method of analysis also presents limitations. Because the variables, both independent and dependent, were mostly categorical, the responses had to be converted before any analysis. As categorical data have no mathematical order, binary variables were assigned to each categorical variable. While collecting categorical data was appropriate for this type of research, not all responses could be appropriately converted into mathematical figures. Hence, there are responses to questions that were omitted from the logistic regression analyses. This type of regression analysis is also subject to limitations, as it is sometimes vulnerable to overconfidence. In other words, the model may report high accuracy levels at predicting outcomes, when in reality the model is not as accurate (Robinson, n.d.).
RESULTS

In this section, my analysis focuses on determining the validity of the sample group, calculating various cross tabulations, and identifying which variables influence likelihood to buy alcohol online through binary logistic regression analyses. I concluded that only 5 of the 13 selected independent variables have statistical significance over influencing the likelihood of a respondent to use e-commerce platforms to buy alcohol.

Prior to any analyses, I omitted 97 of the total 369 responses collected, which is roughly 26% of the sample. The omitted responses were from respondents who don’t drink alcohol. Thus, the analyses reported below only includes responses from alcohol drinkers. The proportion of non-alcohol drinks of the sample is similar to the US population. In 2017, a study estimated that 74% of adults (21 years or older) in the US drink alcohol, and 26% of the US population do not (Grant et al., 2017, p. 914). I also manually changed two entries for the first question, which asked respondents to enter their age in years. The responses were changed for the following reasons:

- One of the respondents answered “1995.” I’m assuming the respondent incorrectly read the question and provided their birth year instead of age. While I do not know this respondent’s actual birth date, I altered this entry to reflect 23 years.

- Another response to this question said “60’s.” Instead of omitting this response, I altered the entry to reflect 65 years as a reasonable estimate of that respondents’ age.
Demographics

Of the 272 responses, I categorized the respondent’s ages into five buckets, as determined by the Kaiser Family Foundation (KFF), an organization for national health and demographic polling and analyses (Kaiser Family Foundation, 2017). Table 4.1 shows the percent of respondents in each age category. To ensure that the age distribution of the sample reflects the United States population, I used 2016 census information from the KFF and compared the percentages in each age bucket of the sample group to the national average, which can also be found in Table 4.1. The distribution of respondent’s ages is somewhat skewed towards older age brackets.

Table 4.1
Age distribution of sample population versus United States population

<table>
<thead>
<tr>
<th>Respondent Age (in years)</th>
<th>Sample Group (n=272)</th>
<th>US Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 - 25</td>
<td>5%</td>
<td>9%</td>
</tr>
<tr>
<td>26 - 34</td>
<td>18%</td>
<td>12%</td>
</tr>
<tr>
<td>35 - 54</td>
<td>35%</td>
<td>26%</td>
</tr>
<tr>
<td>55 - 64</td>
<td>24%</td>
<td>13%</td>
</tr>
<tr>
<td>65 +</td>
<td>19%</td>
<td>15%</td>
</tr>
</tbody>
</table>

Note. 2016 US population age distribution were pulled from Kaiser Family Foundation (2017)

Other demographic variables were relatively reflective of United States census data. Of the responses, 51% (n=139) were female and 49% (n=133) were male, which exactly matches national gender statistics (Kaiser Family Foundation, 2017). Responses of education attainment levels also followed a relatively similar distribution to the United States population where on average, respondents obtained some college or a bachelor’s
degree. Full education demographics can be seen in Table 4.2. (Note. The United States Census Bureau estimated in 2017 the percentage Americans with either No College and Some College in the same category.)

Table 4.2
*Distribution of sample population’s education attainment level*

<table>
<thead>
<tr>
<th>Education Attainment Level</th>
<th>Sample Total (n=272)</th>
<th>US Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>No College</td>
<td>6%</td>
<td>48%</td>
</tr>
<tr>
<td>Some College</td>
<td>34%</td>
<td></td>
</tr>
<tr>
<td>Bachelors</td>
<td>31%</td>
<td>33%</td>
</tr>
<tr>
<td>Graduate</td>
<td>24%</td>
<td>16%</td>
</tr>
<tr>
<td>PhD</td>
<td>5%</td>
<td>3%</td>
</tr>
</tbody>
</table>

*Note. 2017 US Population data were pulled from the United States Census Bureau (United States Department of Commerce, 2017). No College and Some College categories for US population were reported as one group.*

Responses for annual income level also reflected national averages. 48% (n=131) of the respondents reported an income level between $0 - $74,999, where the United States median income level in 2016 was approximately $60,000 annually (Kaiser Family Foundation, 2017).

**Cross Tabulations**

Prior to performing logistic regressions, I decided to uncover preliminary data trends by creating cross tabulation tables with variables and the likelihood to buy alcohol online. These variables were chosen for the crosstab analyses given the research discussed in the literature review.
Age. While research on consumer behavior online suggests that younger people are more likely to buy online (Dillon and Reif, 2004, p. 12), the sample population expressed similar levels of intent to buy alcohol online across each age bracket. Figure 4.1 shows a breakdown of the likelihood indicators of buying alcohol online by each age bracket. For the most part, the distribution of likelihood indicators was homogenous across each age bracket. Although, the oldest age bracket (65+ years) indicated slightly higher levels never intending to use e-commerce to buy alcohol and lower levels of likelihood to use e-commerce to buy alcohol. Surprisingly, the youngest age bracket (19-25 years) was distributed in similar proportions as the oldest bracket. Overall, because the distribution of likelihood indicators is similar across the age brackets, the crosstab analysis suggests that age does not affect the likelihood to buy alcohol online.

![Figure 4.1. Distribution of likelihood indicators (i.e. Never buy alcohol online, Not likely, Somewhat likely, and Likely) in each age bracket](image-url)

26
Gender. Research discussed in the literature review suggested that gender has little effect on buying behaviors online. Figure 4.2 examines likelihood indicators by gender of the sample population. The distribution of each likelihood indicator is fairly homogenous between males and females, thus supporting research discussed in the literature review that gender does not affect a respondent’s intent to buy alcohol online.

![Figure 4.2. Distribution of likelihood indicators (i.e. Never buy alcohol online, Not likely, Somewhat likely, and Likely) by gender](image)

Education Attainment Level. Another demographic variable analyzed in cross tabs was the respondents’ education attainment level. Figure 4.3 breakdown the analysis of each education attainment level and likelihood to buy alcohol online. The distribution of the likelihood indicators was somewhat homogenous with the exception of trends in the No College and PhD categories. The No College bar shows that 0% of respondents in
the category are likely to buy online. However, only 6% (n=16) of the total sample population belongs in the No College category, so the data suggests that this data trend is comparatively insignificant. A separate unusual trend in the PhD category indicates that 77% of respondents who have gotten their PhD’s are not likely to buy alcohol online. However, respondents in the PhD category only represent 5% of the total sample (n=13). While the data suggests slight differences in likelihood across education attainment levels, further analysis is needed with larger sample sizes to make a reasonable conclusion.

![Distribution of likelihood indicators by education attainment level](image)

*Figure 4.3. Distribution of likelihood indicators (i.e. Never buy alcohol online, Not likely, Somewhat likely, and Likely) by education attainment level*

**Employment Status.** The analysis of likelihood by employment status also indicates little significance. Figure 4.4 examines the distribution of likelihood indicators by employment status. Similar to other demographic variables, homogeneity of likelihood
indicators across employment categories suggest that employment status doesn’t affect a respondent’s likelihood to buy alcohol online.

![Figure 4.4](image)

**Figure 4.4.** Distribution of likelihood indicators (i.e. Never buy alcohol online, Not likely, Somewhat likely, and Likely) by employment status

**Income.** A crosstab analysis of likelihood indicators by income level also implies no significant correlation. Data in Table 4.3 suggests that respondents in each income bracket is fairly equally split between feeling not likely (Never and Not likely) or likely (Somewhat likely or Likely) to use e-commerce to buy alcohol. However, the two highest income brackets (represent annual income $175,000+) are slightly skewed towards feeling more likely to use e-commerce for alcohol purchases.
Table 4.3

_Distribution of likelihood indicators by income level_

<table>
<thead>
<tr>
<th>Likelihood Indication</th>
<th>Annual Income Level (in USD)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$0-$9,999 (n=12)</td>
<td>$10,000-$24,999 (n=29)</td>
<td>$25,000-$49,999 (n=50)</td>
<td>$50,000-$74,999 (n=39)</td>
<td>$75,000-$99,999 (n=41)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>0% 42%</td>
<td>24% 55%</td>
<td>28% 60%</td>
<td>13% 44%</td>
<td>10% 49%</td>
<td></td>
</tr>
<tr>
<td>Not likely</td>
<td></td>
<td>31% 32%</td>
<td>31% 44%</td>
<td>39% 59%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somewhat likely</td>
<td>42% 58%</td>
<td>28% 45%</td>
<td>26% 40%</td>
<td>49% 56%</td>
<td>44% 51%</td>
<td></td>
</tr>
<tr>
<td>Likely</td>
<td>17% 58%</td>
<td>17% 45%</td>
<td>14% 40%</td>
<td>8% 56%</td>
<td>7% 51%</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The likelihood indicators are separated by dotted line for further analysis. Respondents who indicated Never or Not likely were grouped together and compared to respondents who indicated either Somewhat likely or Likely.

Table 4.3

_Distribution of likelihood indicators by income level (continued)_

<table>
<thead>
<tr>
<th>Likelihood Indication</th>
<th>Annual Income Level (in USD)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$100,000-$124,999 (n=41)</td>
<td>$125,000-$149,999 (n=21)</td>
<td>$150,000-$174,999 (n=20)</td>
<td>$175,000-$199,999 (n=6)</td>
<td>$200,000+ (n=13)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>12% 59%</td>
<td>14% 48%</td>
<td>15% 55%</td>
<td>17% 33%</td>
<td>8% 31%</td>
<td></td>
</tr>
<tr>
<td>Not likely</td>
<td>46% 33%</td>
<td>40% 35%</td>
<td>40% 50%</td>
<td>17% 67%</td>
<td>25% 54%</td>
<td></td>
</tr>
<tr>
<td>Somewhat likely</td>
<td>34% 41%</td>
<td>33% 52%</td>
<td>35% 45%</td>
<td>50% 67%</td>
<td>54% 69%</td>
<td></td>
</tr>
<tr>
<td>Likely</td>
<td>7% 19%</td>
<td>19% 45%</td>
<td>10% 45%</td>
<td>17% 67%</td>
<td>15% 69%</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The likelihood indicators are separated by dotted line for further analysis. Respondents who indicated Never or Not likely were grouped together and compared to respondents who indicated either Somewhat likely or Likely.

**Shipping Preferences.** Upon researching a variety of e-commerce platforms that sell alcohol, companies often require a minimum order total or a shipping fee, so I wanted to understand how respondent’s shipping preferences related to their likelihood to use e-commerce to purchase alcohol. Figure 4.5 dissects each likelihood indicator by current e-commerce shipping preferences. Since the distribution of the sample is skewed
by 2 respondents in the Always Pay Extra category, a larger sample is necessary to ascertain meaningful results from this data.

Figure 4.5. Distribution of likelihood indicators (i.e. Never buy alcohol online, Not likely, Somewhat likely, and Likely) by e-commerce shipping preferences

**Reasons for Buying Online.** I was also interested to see if the reasons people buy items online had any relationship with their likelihood to specifically buy alcohol online. Figure 4.6 investigates what respondents reported as the main reason they buy items online and how likely they are to buy alcohol online. Since likelihood indicators are relatively homogenous, the data suggests that there isn’t a strong relationship between why a respondent buys online with their likelihood of buying alcohol online.
Figure 4.6. Distribution of likelihood indicators (i.e. Never buy alcohol online, Not likely, Somewhat likely, and Likely) by reason respondent uses e-commerce

**Shipping Expectation for Alcohol Online.** Aside from understanding which variables are influential on the likelihood to buy alcohol online, I also wanted to gauge respondent’s knowledge on the alcohol e-commerce industry. Many platforms (i.e. Drizly, Minibar, and Amazon PrimeNow) offer alcohol deliveries within 1-2 hours. However, Table 4.4 suggests that only 6% (n=16) of the sample population actually expect online alcohol orders to arrive in under 2 hours.
Table 4.4

Sample population’s shipping expectations for alcohol orders online

<table>
<thead>
<tr>
<th>Alcohol E-commerce Shipping Expectation</th>
<th>Percentage of Sample Total (n=272)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within one hour</td>
<td>2%</td>
</tr>
<tr>
<td>1-2 hours</td>
<td>4%</td>
</tr>
<tr>
<td>Same-day</td>
<td>8%</td>
</tr>
<tr>
<td>Next-day</td>
<td>9%</td>
</tr>
<tr>
<td>1-2 days</td>
<td>25%</td>
</tr>
<tr>
<td>2-5 days</td>
<td>34%</td>
</tr>
<tr>
<td>5-7 days</td>
<td>14%</td>
</tr>
<tr>
<td>7+ days</td>
<td>4%</td>
</tr>
</tbody>
</table>

What Type of Beverage Will Be Ordered Online. I was also interested to know which type of alcoholic beverage respondents are most likely to buy online. Figure 4.7 displays the distribution of what the sample usually drink most often versus the distribution of the United States population. Respondents were equally split across each type of alcoholic beverage consumed most often.

![Chart showing distribution of beverage consumption]

Figure 4.7 Distribution of the type of alcoholic beverage the sample population consumes most versus the United States population
type of beverage, however, a Gallup poll indicates that adults in the United States are skewed towards preferring beer over wine and liquor (McCarthy, 2017). Respondents were also asked to indicate which type of alcoholic beverage they are more likely to buy online. While respondents were equally split across each type of beverage consumed most often, responses indicated that the sample was skewed towards feeling more likely to buy Liquor (42%) and Wine (41%) over Beer (17%) when buying alcohol online.

**Logistic Regression Analyses**

To identify which variables influence the likelihood of buying alcohol online, I conducted a logistic regression analysis with all 13 independent variables on the likelihood variable. This type of regression model was used because it allowed me to create a model tailored to predict binary outcomes. In my analyses, the binary outcome is whether a respondent is likely (indicated somewhat likely or likely) or not likely (indicated never or not likely) to use e-commerce to buy alcohol online. Table 4.5 shows the results from the original regression. These results indicate that only 5 of the 13 variables are statistically significant in predicting a respondent’s likelihood of buying alcohol online. Statistical significance was determined if the $p$-value was less than or equal to 0.10. The statistically significant variables in the original model are: Frequent Shoppers, Frequent Alcohol Buyers, Longer Delivery Expectation, Consume Immediately, and Liquor. To determine whether certain variables included in the original model may be causing other significant variables to show up as insignificant, I used a regression reduction process to identify additional significant variables.
Following the statistical method of reducing regression models, I identified the variable with the highest p-value from the original model: Part/Full-Time Employed. A second logistic regression analysis was performed without the Part/Full-Time Employed variable. From the second regression output, the next variable with the highest p-value was identified and removed for a third logistic analysis. This process was repeated eight times until only statistically significant variables remained. See Table 4.6 to follow which variable was removed in each iteration of the regression analysis reduction process.
Table 4.6
*List of logistic regression iterations by variable removed*

<table>
<thead>
<tr>
<th>Logistic Regression</th>
<th>Variable Removed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>All 13 variables were included</td>
</tr>
<tr>
<td>First Iteration</td>
<td>Part/Full-Time Employed</td>
</tr>
<tr>
<td>Second Iteration</td>
<td>Frequent Drinkers</td>
</tr>
<tr>
<td>Third Iteration</td>
<td>Bach/Grad/PhD</td>
</tr>
<tr>
<td>Fourth Iteration</td>
<td>Wine</td>
</tr>
<tr>
<td>Fifth Iteration</td>
<td>Age</td>
</tr>
<tr>
<td>Sixth Iteration</td>
<td>Female</td>
</tr>
<tr>
<td>Seventh Iteration</td>
<td>Store</td>
</tr>
<tr>
<td>Eighth Iteration/Final Reduced Model</td>
<td>Income</td>
</tr>
</tbody>
</table>

On the eighth iteration, the five remaining variables with *p*-values less than 0.10 were Frequent Shoppers, Frequent Alcohol Buyers, Longer Delivery Expectation, Consume Immediately, and Liquor. This data suggests that only 5 of the selected 13 variables have statistical significance on the likelihood to buy alcohol online. Table 4.7 shows the logistic regression output of the final reduced model.
Table 4.7

Logistic regression results from final reduced model

<table>
<thead>
<tr>
<th>Variable</th>
<th>coeff $b$</th>
<th>p-value</th>
<th>exp($b$)</th>
<th>90% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequent Shoppers</td>
<td>1.134</td>
<td><strong>0.000</strong></td>
<td>3.108</td>
<td>1.845 - 5.234</td>
</tr>
<tr>
<td>Longer Delivery Expectation</td>
<td>-0.856</td>
<td><strong>0.004</strong></td>
<td>0.425</td>
<td>0.259 - 0.697</td>
</tr>
<tr>
<td>Liquor</td>
<td>0.493</td>
<td><strong>0.076</strong></td>
<td>1.638</td>
<td>1.036 - 2.589</td>
</tr>
<tr>
<td>Consume Immediately</td>
<td>0.654</td>
<td><strong>0.019</strong></td>
<td>1.924</td>
<td>1.215 - 3.044</td>
</tr>
<tr>
<td>Frequent Alcohol Buyers</td>
<td>-0.792</td>
<td><strong>0.012</strong></td>
<td>0.453</td>
<td>0.270 - 0.759</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval; $LL =$ lower limit; $UL =$ upper limit. Statistically significant variables are indicated by $p$-values less than or equal than 0.10 and appear in boldface.

**Frequent Shoppers Variable.** There are several trends from the regression output that can be used to interpret the significance of how often a respondent buys items online on their likelihood to use e-commerce platforms to buy alcohol. For this variable, the output describes the likelihood of someone buying alcohol online, given the respondent buys items online more than 4 times a month. The $\text{coeff}(b)$ for this variable is a positive number. This suggests that people who buy items online more than 4 times a month are more likely to use e-commerce platforms to buy alcohol than those who buy items online less than 4 times a month. The $\text{exp}(b)$ value, which is the exponentiation of $\text{coeff}(b)$, indicates the odds-ratio of the variable. So, the odds of a respondent indicating they are likely to use e-commerce to buy alcohol are 3.1 times higher when that person buys items online more than 4 times a month. Therefore, the data suggests that frequency a person currently uses e-commerce platforms positively influences their likelihood to specifically use e-commerce to buy alcohol.

**Frequent Alcohol Buyers Variable.** Another significant factor influencing likelihood to buy alcohol online is how often respondents currently buy alcohol. This
variable indicates respondents who buy alcohol more than one time per week. The \( \text{coeff}(b) \) value is negative, which suggests that people who buy alcohol more than once a week are less likely to use e-commerce to buy alcohol. The \( \text{exp}(b) \) value indicates that the odds that a respondent is likely to buy alcohol online is 2.22 times higher when the respondent purchases alcohol less than once a week. This data suggests that people who purchase alcohol less frequently are more likely to use e-commerce platforms for alcohol purchases.

**Longer Delivery Expectation Variable.** Another significant variable that influences the likelihood of buying alcohol online is how long respondents expect their online orders to arrive. Because the \( \text{coeff}(b) \) is negative, the data suggests that if a person expects their online orders to arrive 3 days or more after placing an order, they are less likely to buy alcohol online. More specifically, the odds of a person indicating they are likely to buy alcohol online are 2.38 times higher when they expect online purchases to arrive in less than 3 days. Hence, the data suggests that the longer respondents expect online orders to arrive, the less likely they are to buy alcohol online.

**Consume Immediately Variable.** Another statistically significant variable influencing the likelihood to use e-commerce platforms for alcohol purchases is how quickly a respondent consumes their alcoholic purchases. Regression output for this variable represents respondents who consume their alcoholic purchase within 2 days. So, the positive coefficient suggests that people who consume their alcoholic beverages within 2 days of purchase are more likely to buy alcohol online. Additionally, the \( \text{exp}(b) \) value suggests that the odds of a respondent indicating likelihood to buy alcohol online is 1.9 times higher when the respondent consumes their alcoholic purchases within 2 days.
**Liquor Variable.** The model suggests that the alcoholic beverage that respondents consume most often also influences their likelihood to buy alcohol online. The Liquor variable indicates if a respondent prefers drinking liquor/spirits/mixed drinks over wine or beer. The \( \text{coeff}(b) \) of the Liquor variable is positive, which indicates that respondents who prefer liquor over beer and wine are more likely to buy alcohol online. An interpretation of the \( \text{exp}(b) \) suggests that the odds of a respondent indicating likelihood to buy alcohol online are 1.6 times higher when the respondent prefers liquor over beer and wine. This result is concurrent with the cross tabulation analysis in Figure 4.7 which showed that 42% of respondents (n=114) indicated they would likely purchase liquor over beer or wine on e-commerce platforms.

**Classification Table.** Another meaningful result from the logistic regression model is the classification table, which is presented in Table 4.8. While the eighth logistic regression only included statistically significant variables, Table 4.8 provides an analysis on how well the logistic regression model fits the data. Meaning, how accurate the model was in predicting the Likelihood variable for each respondent. Table 4.8 compares the number of respondents who actually indicated likelihood to buy alcohol online with the number that the model predicted. So, out of the respondents who actually indicated likelihood to buy alcohol online (n=133) in the survey, the model only predicted that 75 of those respondents would indicate likelihood. So, the accuracy of the model’s prediction of likely respondents 56%.

A similar interpretation can be made looking at the Not Likely column. Out of the people in the sample that indicated they were not likely or would never buy alcohol online (n=139), the model only predicted that 98 of those respondents would indicate not
likely. Hence, the accuracy of the model to predict if respondents are not likely to buy alcohol online is 71%. By averaging the model’s accuracy levels of Likely (56%) and Not Likely (71%), we can ascertain that the overall accuracy of the model in 64%.

Table 4.8
*Classification table from final reduced logistic regression model*

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Likely</td>
<td>Not Likely</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likely</td>
<td>75</td>
<td>41</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Not Likely</td>
<td>58</td>
<td>98</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>133</td>
<td>139</td>
<td>272</td>
<td></td>
</tr>
</tbody>
</table>

In addition to calculating the accuracy levels, a further interpretation of the goodness of fit of the logistic regression model can be assessed from Table 4.8. Imagine a scenario where this logistic regression model is unavailable and only the raw data can be used to predict who are people that are likely or not likely to use e-commerce to buy alcohol online. Since more people in the sample indicated that they are not likely to use e-commerce to buy online (n=139; 51%), a reasonable supposition would be to assume that everyone is not likely to use e-commerce to buy alcohol. Using this method to assign likely or not likely to individuals, the raw data suggests that the supposition will be correct for only 51% of individuals. However, if the results from the logistic regression are available, the accuracy of predicting those who are likely or not likely individuals will
increase to 64%. This suggests that the logistic regression model is more beneficial than using raw data to predict those who are likely to use e-commerce to buy alcohol.
IMPLICATIONS

In the following section, I’ve highlighted key implications from trends found in cross tabulation and logistic regression analyses. These implications aim to rationalize how this research can be applied to improve business decisions for companies that sell alcohol online or companies that engage with businesses that do, like manufacturers, wholesalers, and retailers.

Implications of Cross Tabulations

While not all survey variables proved to be statistically significant in the logistic regression analysis, useful insights were found using simpler statistical measurements. Reasonable assumptions and takeaways from the cross tabulation analyses are discussed below.

Demographic Variables. Creating crosstabs with demographic variables and the likelihood to buy alcohol online suggested that demographic categories have little impact on a respondent’s likelihood to use e-commerce to buy alcohol. Distributions of likelihood indicators by age, gender, education attainment level, employment status, and income level were relatively homogenous, meaning people indicated similar likelihood levels across different demographic categories. These results are consistent with results found in the logistic regression model; through the process of model reduction, where only statistically significant variables remain, each demographic variable was removed.
Because demographic categories don’t affect the likelihood to buy alcohol online, the data implies that e-commerce platforms that sell alcohol may not be able to solely use demographic information to segment and target likely customers.

**Shipping Preferences.** Respondents were asked to indicate what their shipping expectations are when they buy items online from the following options: (1) I utilize free shipping, (2) Sometimes I’ll pay extra to have my order delivered faster, or (3) I always pay extra to have my order delivered faster. Of these shipping preference options, respondents’ likelihood indication was relatively homogenous. The data suggests a respondents’ shipping preference does not influence their likelihood to buy alcohol online. Hence, offering promotions on shipping may not be an effective targeting approach to increase traffic to e-commerce platforms that sell alcohol.

**Reasons for Buying Online.** Responses indicating why people use e-commerce resulted in uniformity across likelihood indicators. People who use e-commerce because (1) the product wasn’t available in-store, (2) buying online saved the time of having to buy in-store, or (3) the product was cheaper online, don’t seem to differ in terms of how likely they feel towards buying alcohol online. Therefore, the data suggests that companies may not be able to use the information to identify people with higher likelihood to buy alcohol online.

**Shipping Expectation for Alcohol Online.** Simple trends on a respondents’ shipping expectation for alcohol online revealed that only 6% (n=16) of the sample group thought alcohol deliveries would take less than 2 hours, where 52% (n=141) of the sample thought it would take longer than 2 days. Since major alcoholic e-commerce platforms offer delivery in 1-2 hours, the results imply that companies should invest in
product awareness campaigns to educate people on how quickly online alcohol orders arrive.

**What Type of Beverage Will Be Ordered Online.** In addition to asking respondents to indicate which alcoholic beverage they consume the most, another question asked respondents to indicate which type of alcoholic beverage they are most likely to buy online. Results indicate that 42% (n=114) of the sample are more likely to buy liquor and 41% (n=112) are more likely to buy wine online. Wholesalers can utilize this information by increasing liquor and wine inventories to accommodate the potential of increased demand on e-commerce platforms. E-commerce companies can also use this information to target promotions towards liquor and wine drinkers to increase traffic. The data may also suggest that e-commerce platforms should delay targeting beer drinkers, since likelihood levels are lower.

**Concerns with Alcoholic E-Commerce Platforms.** Aside from attempting to understand which factors make consumers more likely to use e-commerce to buy alcohol, the last question of the survey asked respondents report concerns with buying alcohol online. 43% of respondents (n=117) seemed most concerned with the requirement to sign for the delivery. This suggests an opportunity for e-commerce platforms to offer specific delivery windows to accommodate the concern of having to sign for the delivery. 28% of respondents (n=76) indicated a concern with how long the delivery would take. An implication of this data indicates that there is an opportunity for e-commerce platforms to increase product awareness to educate people on how quickly alcohol can be delivered. When prompted to write-in additional concerns, respondents also questioned the quality of products once it’s delivered. The popularity of this concern amongst respondents
indicates an opportunity for e-commerce companies that sell alcohol to advertise that online orders will refunded/replaced if delivered broken.

**Implications of Logistic Regression Results**

Of the original 13 selected variables included in the logistic regression, only five variables indicated statistical significance. Below, I’ve discussed several implications for each significant variable.

**Frequent Shoppers Variable.** As discussed in the results section, the logistic regression model indicated that the more people currently buy items online, the more likely they are to use alcoholic e-commerce platforms. These results are consistent with findings from Dillon & Reif (2004). Results from the 2004 study indicate that the more experienced internet users reflected a positive attitude towards specifically buying textbooks online. The same conclusion can be made using the logistic regression output. From a marketing standpoint, it seems that companies selling alcohol through e-commerce can utilize this conclusion to identify target markets.

**Frequent Alcohol Buyers Variable.** Another variable that indicated statistical significance in the logistic regression model was how often a respondent buys alcohol. Results imply that respondents who drink alcohol less than twice a week are more likely to buy alcohol online. This conclusion may suggest that, because alcoholic e-commerce platforms have expensive shipping costs, people who buy alcohol more frequently don’t want to incur shipping costs each time they buy alcohol. However, companies in the alcoholic e-commerce industry can use this information to target promotions towards people who are unlikely to use e-commerce to buy alcohol. For example, to encourage
people who buy alcohol frequently (more than twice a week), who are currently unlikely to buy alcohol online, e-commerce platforms can offer an incentive program with a discount on every 10th order.

**Longer Delivery Expectation Variable.** Results from the logistic regression output suggest that the likelihood to buy alcohol online was negatively influenced if a respondent usually expects online orders to take more than 2 days to deliver. This conclusion indicates that respondents are less likely to utilize e-commerce platforms to buy alcohol if they think regular online orders take longer to arrive. E-commerce companies can utilize this information to prioritize marketing campaigns that advertise how quickly online alcohol deliveries arrive compared to delivery times of other products.

**Consume Immediately Variable.** Data from the logistic regression output also indicated that how quickly a respondent consumes their alcoholic products after purchase influences their likelihood to buy online. The model suggests that if a respondent consumes their purchase within 2 days, they are more likely to buy alcohol online. This information can be utilized in conjunction with other trends found in cross tabulation analyses. Another survey question asked respondents to indicate how long they expect online alcohol orders to arrive. Only 22% of the sample (n=60) indicated that they expect their online alcohol order to arrive within 2 days. So, the regression results suggest that (1) people who consume alcohol purchases quickly are more likely to buy online, and the crosstab data suggests that (2) respondents don’t expect alcohol orders to arrive quickly. These results suggest an investment in marketing campaigns to advertise how quickly alcohol orders arrive may improve overall traffic to e-commerce platforms.
**Liquor Variable.** Another statistically significant variable from the logistic regression output indicates that respondents who prefer drinking liquor/spirits/mixed drinks over beer and wine are more likely to use e-commerce to buy alcohol. This result corresponds with responses to another question in the survey which asked respondents to indicate which type of product they are more inclined to buy online. Roughly 42% of the sample population (n=114) indicated that they were more likely to buy liquor over beer or wine on e-commerce platforms. Companies in the alcohol e-commerce industry can utilize this information to deliver targeted promotions towards liquor drinkers, since they indicated higher likelihood to use e-commerce to purchase alcohol. Wholesalers and retailers can potentially use this information to inform product assortment decisions. For example, wholesalers may decide to stock more liquor products if they have significant e-commerce activity.
CONCLUSION

With any new and emerging industry, researchers and marketers attempt to make sense of consumer behavior to better inform business decisions. While the alcohol industry has existed for centuries and the e-commerce industry since the 1990’s, e-commerce platforms that sell alcohol are just beginning to emerge.

The purpose of this research was to uncover answers about which factors influence the likelihood of someone to use e-commerce to specifically purchase alcohol. By selecting 13 independent variables that represent demographic characteristics, current e-commerce behavior, and alcohol purchase habits, this research aimed to clearly identify which characteristics make a person more likely to buy alcohol online.

Results from cross tabulation and logistic regression analyses conclude that only five variables significantly influence the likelihood to buy alcohol online: (1) the frequency a respondent currently buys items online, (2) how often a respondent buys alcohol, (3) how long a respondent expects regular online orders to arrive, (4) how many days after purchase a respondent waits to consume an alcoholic purchase, and (5) the type of alcoholic beverage respondents consume the most. After analyzing the marginal effects of each significant variable, I was able to make conclusions about how companies selling alcohol online can make more meaningful marketing, promotional, and inventory decisions.

While I was able to uncover some influential factors on the likelihood to buy alcohol online, there are many questions for further study. Since shipping costs for
alcohol orders on e-commerce platforms are generally more expensive than other retail e-commerce platforms, marketers may want to perform research on the price sensitivity of shipping costs. Another pertinent question is how a larger sample size might alter the results presented in this thesis. Since some variables were skewed, perhaps a larger sample population may alter which variables are statistically significant. My survey was also only distributed to people in the United States. So, conducting research on the likelihood to use e-commerce to buy alcohol in international markets may yield different results and imply altered call-to-actions for companies. Finally, as e-commerce continues to grow in the United States and consumer behavior continues to change, opportunities to re-evaluate this research will present themselves in the coming years to discover the most up-to-date influential factors on the likelihood to use e-commerce to buy alcohol.
APPENDIX

A: Survey Consent Form

The purpose of this research study is to see how consumers will behave when buying alcohol online. The research is designed to benefit society by gaining new knowledge. There is little chance that you will benefit from being in this research study.

Being in a research study is completely voluntary. You can choose not to be in this research study. You can also say yes now and change your mind later.

If you agree to take part in this research, you will be asked to answer the following questions in this survey. Your participation in this study will take about 5 minutes. We expect that 500 people will take part in this research study.

You can choose not to answer any question you do not wish to answer. You can also choose to stop taking the survey at any time. **You must be at least 21 years old to participate. If you are younger than 21 years old, please stop now.**

To protect your identity as a research subject, no identifiable information will be collected, the research data will not be stored with your name, the researcher(s) will not share your information with anyone. In any publication about this research, your name or other private information will not be used.

If you have any questions about this research, please contact the Investigator by emailing Amanda_Schroeder@kenan-flagler.unc.edu. All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject, or if you would like to obtain information or offer input, you may contact the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.
B: Survey

Note: The * indicates that an answer is required.

* 1. How old are you (in years)?

* 2. What is your gender?
   - Male
   - Female

* 3. What is your education level?
   - No College
   - Some College
   - Bachelors
   - Graduate
   - PhD

* 4. What is your employment status?
   - Student
   - Unemployed
   - Part-time employed
   - Full-time employed
5. What is your income level per year?

- $0-$9,999
- $10,000-$24,999
- $25,000-$49,999
- $50,000-$74,999
- $75,000-$99,999
- $100,000-$124,999
- $125,000-$149,999
- $150,000-$174,999
- $175,000-$199,999
- $200,000+

6. How often do you purchase items online?

- Once a month
- 2-4 times a month
- 4-10 times a month
- 11+ times a month

7. For your most recent online purchase, why did you choose to buy online as opposed to in-store?

- You went in-store and it wasn’t available
- Buying online saved you the time of having to buy in-store
- Product was cheaper online

8. When you buy online, how quickly do you expect your order to arrive?

- Next day
- 2 days
- 3-5 days
- 7 days
- 8+ days
9. When you buy online, do you pay for shipping when there is an option for free shipping if it means you can get your order faster?

- No, I utilize free shipping
- Sometimes, I'll pay extra to have my order delivered faster
- Yes, I always pay extra to have my order delivered faster

10. How often do you consume alcohol?

- Never
- Less than once per week
- 1-2 times a week
- 2-4 times a week
- 4-7 times a week
- More than 7 times a week

11. What alcoholic beverages do you consume the most?

- I do not consume alcoholic beverages
- Beer
- Wine
- Liquor/Spirits/Mixed Drinks

12. Where do you most often purchase alcohol?

- I do not purchase alcohol
- At grocery stores/liquor stores
- At convenient stores/gas stations
- At restaurants/bars
- Online
* 13. How quickly after you purchase alcohol do you begin consumption?
   - Not applicable (I do not purchase alcohol)
   - Immediately (Within 24 hours)
   - Within 1-2 days
   - Within 7 days
   - Beyond 7 days

* 14. How often do you purchase alcohol at grocery stores, convenient stores, and/or liquor stores?
   - Not applicable (I do not purchase alcohol)
   - Once a week
   - 1-2 times a week
   - 2-5 times a week
   - 6+ times a week

* 15. Did you know you could purchase alcohol online?
   - Yes
   - No

* 16. How often do you purchase alcohol online?
   - Never
   - Less than once a month
   - 1-2 times a month
   - 2-4 times a month
   - Once a week
   - 1-2 times a week
   - 2-5 times a week
   - 6+ times a week

* 17. Given the option to purchase alcohol online, how likely are you to buy alcohol online?
   - Not applicable (I do not purchase alcohol)
   - I would never buy alcohol online
   - Not likely
   - Somewhat likely
   - Likely
18. If you purchased alcohol online, how quickly would you expect your order to arrive?

- Not applicable (I would not purchase alcohol online)
- Within one hour
- 1-2 hours
- Same-day
- Next-day
- 1-2 days
- 2-5 days
- 5-7 days
- 7+ days

19. If you were to purchase alcohol online, what are you most likely to buy?

- Not applicable (I would not purchase alcohol)
- Beer
- Wine
- Liquor/Spirits

As of February 2018, alcohol deliveries require that an adult 21 years or older be present at the time to sign for the delivery, and cannot be left at a residence without a signature.

20. Now knowing this information, how likely are you to buy alcohol online?

- Not applicable (I do not purchase alcohol)
- Somewhat likely
- I would never buy alcohol online
- Likely
- Not likely
21. What might prevent you from buying alcohol online? Check all that apply.

- Delivery time is too long
- Need to sign for delivery
- Nervous about supplying credit/debit card info
- Other (please specify)
REFERENCES


Hui, T., & Wan, D. (2009). Who are the online grocers?. Service Industries Journal, 29(11), 1479-1489. doi:10.1080/02642060902793334

Johnson, K. (2017) E-commerce industry spotlight. Retrieved from Brigham Young University, Marriott Student Review website:
Kaiser Family Foundation. (2017). *2016 Median annual household income.* [Table]. Retrieved from https://www.kff.org/other/state-indicator/median-annual-income/?currentTimeframe=0&sortModel=%7B%22colId%22:%22%22Location%22,%22sort%22:%22asc%22%7D

Kaiser Family Foundation. (2017). *2016 Population distribution by age.* [Table]. Retrieved from https://www.kff.org/other/state-indicator/distribution-by-age/?currentTimeframe=0&sortModel=%7B%22colId%22:%22%22Location%22,%22sort%22:%22asc%22%7D


