E-Commerce Profitability & Consumer Purchase Behavior

By

Kashish Kumar

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Undergraduate College

Leonard N. Stern School of Business

New York University

May 2016

Professor Marti G. Subrahmanyam
Faculty Advisor

Professor Daria Dzyabura
Thesis Advisor

Table of Contents

Abstract

Consumers often evaluate alternatives differently based on whether they are performing the search online, or offline physically in a store. Information is available more systematically online for consumers to evaluate thus enabling them to rate each and every feature of the product. However, certain attributes (such as size) are difficult to judge without the physical inspection of the product and consumers clarify their preferences only after viewing the product offline. In this paper we attempt to understand how preferences change when evaluating products online versus offline. Specifically, we study the change in utility derived from the various attributes of a product and the relationship between the utility to a consumer from these attributes online and offline. Additionally, we analyze the impact of changing preference on the change in willingness to pay for each of the product attributes, and change in willingness to pay for features when it has been evaluated both online and offline. The change in willingness to pay demonstrates how consumers might over-value or under-value features based on their mode of evaluation. Consumers often use the non-compensatory decision making process when making purchase decisions, and the ability to filter products online impacts which products are finally included in the consumer's choice set. In this paper we also attempt to understand if a consumer is more non-compensatory while evaluating products online or offline.

E-Commerce Profitability & Consumer Purchase Behavior

Introduction

In the past decade e-commerce firms have gone from their humble beginnings to becoming a threat to established, behemoth retail chains like Walmart. The rapid growth in internet sales has served to provide consumers with an alternate medium to purchase goods they would traditionally buy at brick and mortar stores. The "touch-and-feel" method of product evaluation is still prevalent, especially for moderate and high-involvement products such as furniture that are purchased relatively in-frequently. Worldwide internet consumer expenditure has gone from 3% in 2006 (as a % of total retail sales) to an estimated 15.5% in 2015 (Cole, 2016). Within the e-commerce space, sales grew more than 20% worldwide in 2014 to almost $840 billion, as e-commerce giants such as Amazon continued to expand to new geographies and traditional retailers also started offering their products virtually. This boom in e-commerce can be seen in the skyrocketing valuations of new firms, the most notable of these being Alibaba which conducted a hugely successful $25 billion IPO in September of 2015, valuing the Chinese internet giant at almost $170 billion (Ben-Shabat, Nilforoushan, Yuen, & Moriarty, 2015). In the United States alone, e-commerce sales grew 14.7% in the fourth quarter of 2015 compared from the fourth quarter of the previous year, totaling $89.1 billion (U.S. Department of Commerce, 2016).

Despite these rapid advancements in e-commerce, many consumers still prefer to visit a physical store to evaluate a product. Researchers have studied and created models to elicit and predict consumer preferences however it is extremely important to understand how consumer preferences change regarding the different attributes of the same product when they assess them online and offline. One of the easiest ways to construct consumer preference models is to conduct a conjoint analysis (Green and Srinivasan, 1978) in which subjects are asked to rate pairs of products in which only one attribute is changed. Using this technique, researchers can determine which attribute a subject prefers in the product and these attributes are then ranked. The value that the subject assigns to each attribute can be calculated as a statistical estimate, which then enable the researcher to map out trade-offs between these attributes.

Most marketing researchers conducted this survey online and then used the results for the offline preference mapping as well, incorrectly assuming that preferences are symmetrical both online and offline. Past research has shown that there is a change in preferences and utilities when evaluating a product online and offline. This phenomenon is very common as consumers usually do not realize how important certain features are to them until they physically see the product. For instance, you might compare and select a desk chair online and decide to visit a physical store to "test" it out. While reading the product description online, you might feel the ability to adjust the height, color and presence of wheels are the most important features, however when you actually see the chair you realize you actually care more about the comfort when you are sitting and whether or not it can recline as the most important features. In fact, you might even reject a chair completely if while evaluating offline you see that one of the features you "weighted" as low utility online is not present. Additionally, it is possible that you were indifferent between the colors black and navy online and would have the same willingness to pay regardless of which color, however when you actually see the chair you realize that the blue is a little too bright for you and you would be willing to pay a premium to acquire the black chair. Using data collected for the study by Dzybura, Jagabathula and Muller, we can test for: 1) changes in the partworth utilities for each attribute – testing if consumers along with a change in preference for features also change their relative preference between pairs of features, 2) effect on willingness to pay based on the relative partworth utilities and 3) test for the changes in the range of the various overall attribute importance to determine if consumers are more/ less non-compensatory offline vs. offline.

Literature Review

In this paper we aim to build on the previous research conducted specifically in the area of Consumer Behavior and consumer purchase decisions making. Our study relies on the accuracy of the conjoint analysis tools developed by Green and Rao (1971) which enables us to elicit targeted responses by subjects about specific features of a product. Over time, this method has been tweaked and improved upon by numerous researchers, thus making the rank, ordering and choice elicitation tasks an effective tool for any marketing research study. Professor of Marketing as Columbia University, Oded Netzer advanced the techniques developed by Rao et al. by identifying and addressing the gaps in traditional preference elicitation and estimation models. Specifically, Netzer focused on addressing three components of preference measurement: 1) the problem that the study intends to address, 2) design and approach of the data collection study and preference measurement task, 3) advanced preference estimation models.

Since then, researchers assumed that the conjoint analysis conducted online was a good estimate of consumer preference when they assess products offline as well. Dzyabura et al. (2016) improved and built a new data fusion model that improved upon the reliability of the online ratings data to predict offline preferences. They demonstrated that large discrepancies existed between the online and offline partworths when consumers evaluated products physically versus online. Since collecting large amounts of data offline is time consuming and extremely expensive, large online data sets can be combined with small offline data sets to better estimate offline preferences (up to 25% improvement). This study borrows extensively and builds on the results found by Dzyabura et al. (2016).

Another area of research that this paper contributes to is the idea of constructed preferences (Dzyabura, The Role of Changing Utility in Product Search, 2013). The focus of the above paper was demonstrating how consumers might be unaware of their preferences when they begin a product search. As they view products, their preferences are formed and their utility functions and attribute weights tend to change. Even though the focus of our study is to measure the change in preferences between online and offline product search, the above paper lays the foundation for our argument about the role of changing consumer behavior.

In this paper we borrow concepts and idea from the above areas of research and aim to further understand the changes in consumer behavior given the rapidly changing retail landscape. The shift in focus to e-commerce makes it extremely important for retailers to accurately understand how consumers might rate and evaluate their offerings if the same is available both online and in a physical store. To the best of our knowledge, no paper has focused on studying this area of research.

Details of the Study

Data Collection

Thanks to the rapid advancements in internet availability across the country, conducting large scale online studies is relatively inexpensive and quick. The large data-set gathered online was then combined with a small subset of the respondents who completed the offline study. The two data sets were then fused together using a statistical data-fusion tool that estimated the offline partworths using two approaches: a hierarchical Bayesian approach and a k-Nearest Neighbors approach to derive the most reliable results. The case study was conducted using

Timbuk2 messenger bags; these bags are relatively expensive, not purchased frequently but something that most people are aware of and have some familiarity with the features, have tons of customization options making it easy to study the different features of the bag, and fits into the category of products that consumers might purchase online while also testing it out in a physical store. Dzyabura, et al. (2016) tested for the following features of the bag to reduce the complexity while ensuring that the data would yield valuable results (Dyzabura, Jagabathula, & Muller, 2015):

- Color (4 options): Black (default), Blue, Reflective and Colorful
- Size (2 options): Small (default – 10 x 19 x 14 in) and Large (12 x 22 x 15 in)
- Price (4 levels): $120 (default), $140, $160 and $180
- Strap pad (2 options): Yes (default) and No
- Water bottle pocket (2 options): Yes (default), No
- Interior compartment (3 configurations): Empty bucket with no dividers, Divider for files and a padded laptop compartment

Note: Price was used as a continuous variable.

122 students from New York University took the tests and adequate steps were taken to ensure incentive compatibility, i.e. each participant would be entered into a raffle with a chance to win a free messenger bag. Additionally, to ensure that the subjects did not randomize their answers, the researchers told the students that the bag would be configured to the winning student's preferences as stated in the study.

Figure 1: Screenshot of the online task to be completed

Figure 2: Offline task set-up in a room separate from the one where the online study was conducted.

Data Manipulation

Estimation of partworths from the ratings data. We used the data collected by Dzyabura et al. (2016) for the purpose of our study. While the previous studies aimed at accurately estimating the offline partworth utilities using the online data, our aim in this study was to analyze the results and test for patterns in changes in behavior online versus offline. Dzyabura et al (2016) proved that the online partworth and offline partworth estimates were different at the 1% significance ($p<0.01$) level.

Calculating overall attribute importance. The non-binary variables such as color, price and interior features were formatted such that the base case for each of them was normalized to the 0. In the table below, the coefficient for "c*olor 1"* in the offline partworths table is *-1.00*; this implies that the first subject receives additional -1.00 utility from switching to color 1 over that base color (black). Binary variables are a little different: size, for instance, in the offline partworths table for the first subject has a coefficient of *0.20;* this implies that for this subject the gain in utility from switching size (from the default – small to large) is 0.20.

Figure 3: Snapshot of the offline and online partworths tables.

Offline Partworths							
color1	color2	color3	size	strap	waterbottle	divider	laptop
-1.00	-0.45	-0.45	0.20	0.44	0.20	0.17	-0.22
0.25	0.24	-0.76	-1.10	0.42	-0.10	0.17	0.22
-0.50	-0.72	-0.22	-0.50	0.69	0.10	1.17	1.30
-0.75	0.47	-0.20	0.10	0.33	-0.10	0.17	-0.18
-0.50	0.99	0.32	-0.60	0.13	0.40	0.33	0.76
-2.00	-1.36	-0.36	-0.80	0.49	0.20	-0.17	-0.97

Online Partworths							
color1	color2	color3	size	strap	waterbottle	divider	laptop
0.00	-0.40	-0.06	0.60	0.65	0.40	-0.50	0.23
0.25	0.06	-0.27	0.10	0.96	0.70	1.33	1.39
-0.25	-1.35	-0.35	0.10	0.26	0.10	0.50	1.23
-0.25	-0.58	-0.25	0.90	0.19	0.30	0.00	0.29
0.00	0.61	0.44	-0.30	0.46	0.30	0.67	1.06
-0.25	-1.60	-0.27	0.10	0.94	-0.30	0.00	-0.69

The overall attribute importance for each of the features of the bag essentially represents if the consumer cares about that particular feature or not. If the presence or absence of a water bottle pocket makes no difference to the consumer, he will have a low overall attribute importance for the variable water bottle. In our study the overall attribute importance is represented by the range of individual partworths for each feature. For the binary features like size, the coefficient of the partworth in the above table is the overall attribute importance. However, for variables with multiple levels, like color, the overall attribute importance is the difference between the maximum weight attributed to a level in that feature and the lowest weight attributed. We took into consideration that the base/ default levels of these features were normalized to 0. For instance, the feature interior has 3 levels – empty bucket with no dividers, divider for files and divider with a padded laptop compartment. The default level was no divider, which had its utility normalized to 0. The overall attribute importance of interior for subject 1 in the online study is thus calculated as (also shown in Figure 4 below):

$$\mathrm{imp}(\mathrm{interior}) = \max\{0, divider, laptop\} - \min\{0, divider, laptop\}$$

$$imp(interior) = 0.23 - (-0.50)$$

$$imp\ (interior) = 0.73$$

Figure 4: Snapshot of the tables containing overall attribute importance weights

Offline Overall Atrribute Importance						Online Overall Atrribute Importance					
color	size	price	strap	waterbottle	interior	color	size	price	strap	waterbottle	interior
1.00	0.20	-0.39	0.44	0.20	0.39	0.40	0.60	-0.37	0.65	0.40	0.73
1.01	-1.10	0.31	0.42	-0.10	0.05	0.52	0.10	-0.34	0.96	0.70	0.06
0.72	-0.50	-1.03	0.69	0.10	0.14	1.35	0.10	-0.61	0.26	0.10	0.73
1.22	0.10	-0.43	0.33	-0.10	0.34	0.58	0.90	-0.27	0.19	0.30	0.29
1.49	-0.60	-0.18	0.13	0.40	0.43	0.61	-0.30	-0.09	0.46	0.30	0.39

Hypothesis and Results

Using the data we gathered, there were a couple of theories/ hypothesis we wanted to test for: 1) Does the relative importance of the different attributes of a product change when a consumer evaluates a product offline versus online? 2) What is the impact on willingness to pay for individual features, different levels within each feature and combinations of different features when evaluating the product offline versus online, 3) Do consumers tend to become more non-compensatory offline versus online?

Hypothesis 1. Prior to conducting the study, we knew that it was common for preferences to change when a consumer evaluates a product offline versus online Dzyabura et al (2016), however we wanted to study how the preference for one feature changes with respect to the others. We conducted a simple correlation analysis on pairs of features for each, online and offline data. The correlation was carried out on the overall attribute importance tables as this would give us a clearer picture on which feature consumers tend to prefer with respect to the other available features. For instance, if the importance of color goes up, what happens to the importance of size? How are consumers making trade-offs between features while evaluating offline versus online?

Figure 5. Pairwise correlation of the various features, offline and online.

Offline						
	color	size	price	strap	waterbottle	interior
color	1.00	0.47	0.32	-0.07	-0.11	0.29
size		1.00	-0.18	0.43	0.33	0.11
price			1.00	-0.15	-0.14	0.00
strap				1.00	0.54	0.06
waterbottle					1.00	-0.09
interior						1.00

Online						
	color	size	price	strap	waterbottle	interior
color	1.00	-0.05	0.29	-0.29	-0.43	-0.13
size		1.00	-0.14	0.34	0.12	-0.20
price			1.00	0.01	0.00	0.06
strap				1.00	0.55	-0.13
waterbottle					1.00	-0.10
interior						1.00

Results 1. We found that for certain pairwise correlation, the relative importance of the feature does change dramatically, i.e. if the increasing importance of one feature tended to result in the increasing importance of another feature online, the opposite happened on the offline study. Take size and interior for instance; in the above table the correlation between size and interior is 0.11, however online the correlation is -0.20. This represents a substantial change (>200%) from offline to online. Here we found that consumers who tended to weigh size as a more and more important factor, also tended to find interior as an important overall decision making factor when the study was conducted offline. The story is quite different online where consumers who viewed size as an increasingly important factor, found the interior features to be of less and less importance. Overall out of the 15 pairs of features tested, more than 25% flipped relative directions, 25% remained practically unchanged (or small changes), and the rest had somewhat significant moves.

Hypothesis 2. Once we were able to establish that along with a shift a preference for features, the relative overall importance of features changes, we were confident that this would impact how much consumers were willing to pay for individual features and how this would change when the study was conducted offline versus online. To calculate this, we used the average relative coefficient of the partworth utilities and the average coefficient of the partworth utility of price. This yielded the coefficient of relative partworth utilities with respect to price, and since price was treated as a continuous variable with increments of $20 in the original study, we multiplied this factor by 20 to get how much a consumer was willing to pay to acquire a particular feature over the base feature.

Figure 6. Change in willingness to pay for levels of features over the base option.

Increase in WTP for pairs of features Offline								
	color1	color2	color3	size	strap	waterbottle	divider	laptop
base	-$79	-$94	-$15	-$41	$34	$23	$69	$116
color1				-$120	-$45	-$56	-$10	$37
color2				-$134	-$60	-$71	-$25	$23
color3				-$55	$19	$8	$54	$102
size					-$7	-$18	$28	$76
price					$14	$3	$49	$96
strap						$56	$102	$150
waterbottle							$91	$139
divider								$185

Increase in WTP for pairs of features Online								
	color1	color2	color3	size	strap	waterbottle	divider	laptop
base	-$28	-$95	-$20	$24	$46	$40	$37	$56
color1				-$4	$18	$12	$8	$27
color2				-$71	-$50	-$55	-$59	-$40
color3				$4	$26	$20	$17	$36
size					$70	$64	$61	$80
price					$26	$20	$17	$36
strap						$86	$82	$101
waterbottle							$77	$96
divider								$92

Note: Willingness to pay was additive. The additional WTP for size and strap was just the sum of their individual change in WTPs.

Results 2. Through this study we were able to make a few observations in terms of willingness to pay:

1) Consumers were willing to pay the most for the base color – black. This trend did not change from offline to online. However, the *color 1* had a positive change in willingness to pay when consumers went from online to offline suggesting that upon physical evaluation of the bags, the subjects tended to dislike *color 1* even more than they did when they saw this option online. In other terms, while evaluating the bags online, the subjects were willing to pay an additional $28 to acquire the color black over *color 1,* however upon physical evaluation they realized they really disliked

color 1 compared to black, and were willing to pay an additional $79 to acquire the color black over *color 1*.

2) We observed one trend in the results table below – the willingness to pay for utilitarian features, such as the presence of a divider or laptop compartment, increased. While evaluating the bags online subjects were willing to pay less to have the additional laptop compartment than they were after having physically evaluated the bags. One explanation for this shift in behavior could be that it is difficult for us to determine the importance of physical, functional features when viewing a product online. Only when we physically assess the product do we start to understand the importance of these features, maybe because we can better visualize what item would go in what bucket when a divider is present.

Figure 7. Change in willingness to pay when going from online to offline.

Change in WTP (online - offline)								
	color1	color2	color3	size	strap	waterbottle	divider	laptop
base	$51	-$2	-$5	$65	$12	$18	-$32	-$61
color1				$115	$63	$68	$18	-$10
color2				$63	$10	$16	-$34	-$63
color3				$60	$7	$13	-$38	-$66
size					$77	$82	$32	$4
strap						$30	-$20	-$48
waterbottle							-$15	-$43
divider								-$93

Note: A positive change implies an increasing distaste towards that level of the feature.

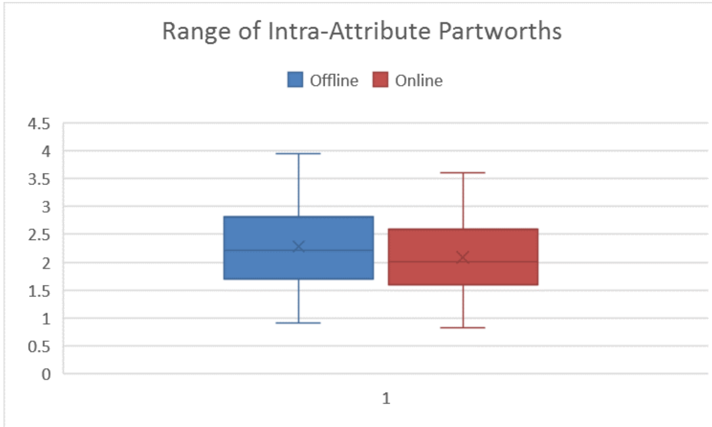
Hypothesis 3. Consumers today are faced with choice paralysis due to the large number of options available in the marketplace that meet their search criteria. To make this decisions making process easier, consumers commonly use non-compensatory decision making techniques (Einhorn, 1970). Non-compensatory decision making involves the use of heuristics; consumers have one or a few *"make or break"* features they would like in the product, which if not present would eliminate the product from their consideration set altogether, no matter how good the other features are. Intuitively we would expect consumers to be more non-compensatory when making decisions online because all the information is readily available, sortable and we can also filter our search results (e.g. Use of maximum price to filter results). However, when we thought about this a bit longer we predicted that consumers might tend to get more non-compensatory while making decisions offline. This is because while assessing a product offline, you might see a feature that you really dislike (or like) and decide that no matter what you will not purchase the product because you never get to evaluating the other features.

To test for non-compensatory behavioral changes, we once again used the overall attribute importance tables. We then conducted two tests here: 1) We calculated the range, variance and other statistics from the overall attribute importance tables for each individual, 2) we plotted the overall attribute importance of each feature for both the online study and the offline study using the box and whisker chart (Figure 10 below).

Figure 8. Analytical statistics on the range of overall attribute importance.

Range of Overall Attribute Weights			
offline		***online***	
Mean	2.282	Mean	2.090
Standard Deviation	0.706	Standard Deviation	0.661
Sample Variance	0.499	Sample Variance	0.437
Range	3.028	Range	2.784
Minimum	0.917	Minimum	0.825
Maximum	3.944	Maximum	3.609
Confidence Level(95.0%)	0.127	Confidence Level(95.0%)	0.119
P(T<=t) 5% Significance level:		0.001263683	

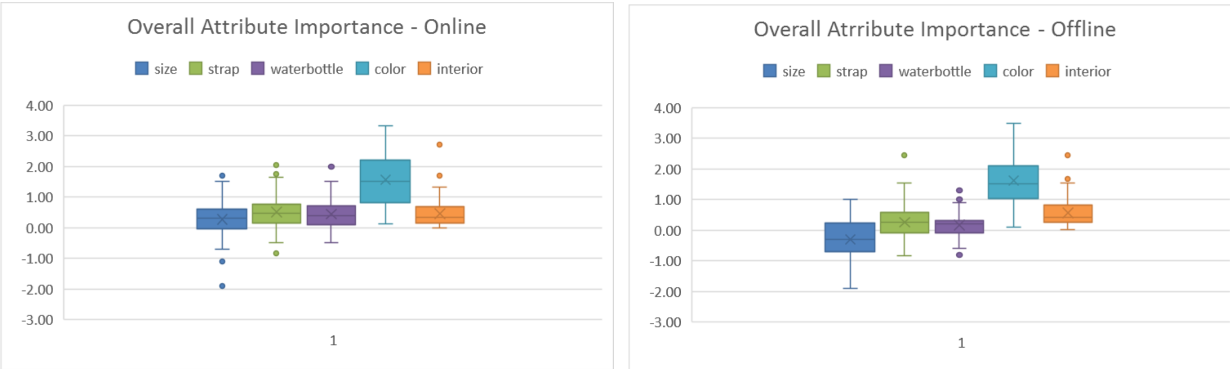
Figure 9. Snapshot of the distribution of values for the range of overall attribute importance



Results 3. We found that consumers do have a tendency to get more non-compensatory when making decisions offline versus online. We tested to see which data set showed the higher range, variance and standard deviation measures. Higher range and variance in overall attribute importance (Figure 8) in the offline study implies that there are some features that the subjects really cared about as they got a lot of partworth utility from either having or giving up those features. This can also be seen in Figure 9.

In Figure 10 we see the breakdown of the ranges of feature level partworth ranges. For the online study we noticed that for each of the attributes, partworth range between the 2nd and 3rd quartile did not contain the value 0. However, for the offline study, 3 out of the 5 features meaningfully contained 0 between the 2nd and 3rd quartiles. This proves that while evaluating offline, consumers tend to care about a few features (color and interior) while all other attributes get low partworth utilities.

Figure 10. Online versus offline overall attribute importance plots



Implications and Conclusion

The rapid shift in the preferred mode of purchasing goods has put the e-commerce firms in a relatively stronger position when compared to traditional retailers. The entire world is headed towards a society that shops by speaking words into their phones, and this means that consumers will develop new ways to make their purchase decisions. Even though the e-commerce market is set to continue to grow for the next decade, we cannot ignore the importance of traditional brick and mortar stores, especially when it comes to high involvement, relatively expensive goods that are purchased infrequently. This is why the fastest growing e-commerce giants like Amazon and Warby Parker are opening up brick and mortar stores.

With the opening up of these brick and mortar stores comes the issue of optimal assortment of products in the offline and online channels (Dzyabura & Jagabathula, Offline Assortment Optimization in the Presence of an Online Channel, 2015). While Dzyabura and Jagabathula discuss the optimal mix of products that should be included in the two channels, it is imperative for firms to understand how consumer behavior changes when they view products both online and offline. This knowledge will enable the firms to determine which features to highlight online, and which features to focus on in the physical stores. Firms can avoid highlighting features that tend to make consumers more non-compensatory offline to ensure that the consumer evaluates all features before making a decision. This will help consumers also make an informed decision and potentially avoid them on missing out on a product which otherwise would have been a great fit but was left out due to the absence of one feature.

Future Work

In this paper we made some simplifying assumptions such as not including search costs, and we ignored the impact of the order in which the consumers rate and evaluate the goods. It is possible that viewing an item offline before looking for it online might include some inherent biases in the search process. The consumer might be more focused on the utilitarian features (like interior pocket in our study) rather than taking into consideration all other features. With this bias, the consumer is likely to filter products online that match what they see physically, thus potentially leaving other products out of their consideration set which might be low on the one feature, but otherwise include the best of all others.

Another interesting theory to test might be the deliberate inclusion of decoy products either online or offline, to further influence the consumer decision making process. This could help firms push consumers into purchasing a product that is high margin/ more profitable. We believe this paper lays the foundation for a new avenue of further research which could be mutually beneficial to both the firms selling products and consumers making the purchase decision.

References

Ben-Shabat, H., Nilforoushan, P., Yuen, C., & Moriarty, M. (2015). *Global Retail E-Commerce Keeps On Climbing*. AT Kearney.

Cole, L. (2016, February 17). *Business 2 Community*. Retrieved from Business 2 Community: http://www.business2community.com/ecommerce/rise-rise-e-commerce-statistics-trends-business-can-capitalize-01456979#TLilozL8RJRkzt7H.97

Dyzabura, D., Jagabathula, S., & Muller, E. (2015). Using Online Preference Measurement to Infer Offline Purchase Behavior. *Available at SSRN 2603264*.

Dzyabura, D. (2013). The Role of Changing Utility in Product Search. *Available Online at SSRN*.

Dzyabura, D., & Jagabathula, S. (2015). Offline Assortment Optimization in the Presence of an Online Channel. *Available at SSRN 2515036*.

Einhorn, H. (1970). The Use of Nonlinear, noncompensatory Models in Decision Making. *Psychological Bulletin*, 73(3), 221.

Green, P. E., & Srinivasan, V. (1990). Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. *Journal of Marketing, American Marketing Association*, 3-19.

U.S. Department of Commerce. (2016). *Quarterly Retail E-Commerce Sales 4th Quarter 2015.* Washington: U.S. Census Bureau.