What Makes ‘em Tick: The Effects of Shopping Mode and Culture on Time and Money Spent Shopping

Written by Hubert W. de Nie (303470)
Supervised by dr. Nuno Camacho

July 27, 2012
# Table of Contents

1. **Introduction: A Shift of Perspective** ........................................................................... 5  
   Objectives .......................................................................................................................... 5  
   Contributions .................................................................................................................... 6  

2. **Conceptual Framework** .............................................................................................. 7  
   Shopping Modes ................................................................................................................ 7  
   Antecedents of a Consumer’s Shopping Mode: The Role of Culture ............................. 11  
      Defining Culture ............................................................................................................. 11  
      Culture as a Predictor .................................................................................................... 13  
   Consequences of Consumer’s Shopping Mode: Purchase Incidence and Time Spent in Store .......................................................................................................................... 14  
      Effects of Shopping Mode on Purchase Incidence ...................................................... 14  
      Effects of Shopping Mode on Time Spent Shopping .................................................. 16  
   The Moderating Effect of Culture on the Relationship between Shopping Mode and Shopper Behavior ................................................................................................................. 17  
      Moderating Effect of Culture on the Relationship between Shopping Mode and Purchase Incidence ............................................................................................................... 17  
      Moderating Effect of Culture on the Relationship between Shopping Mode and Time Spent Shopping ............................................................................................................. 19  
      Overview of Hypotheses ............................................................................................... 20  

3. **Methodology** .............................................................................................................. 20  
   Data .................................................................................................................................. 20  
   Variable Operationalization .............................................................................................. 24  
   Econometric Model .......................................................................................................... 24  
   Data Cleaning ................................................................................................................... 25  

4. **Results** ....................................................................................................................... 26
Breusch-Pagan and Koenker Test ................................................................. 50
Histogram of Residuals Model ................................................................. 50
7. Bibliography ............................................................................................... 51
1. **Introduction: A Shift of Perspective**

Near the end of the 20\textsuperscript{th} century, we have witnessed great changes in the way we live and interact with our surroundings. The rise of the Internet, mobile telephones, and increased ease and affordability of world travel have helped tear down many of the borders which previously defined and distinguished our nations. We are all familiar with this phenomenon, it’s called globalization.

One of the most clearly observable aspects of globalization is the convergence of consumer behavior due to decreasing inequality gaps. Developing countries such as Brazil, China and India grew at unimaginable rates and soon contributed to the ever growing middle class population. With cellphones, Western music channels and the Internet hitting every middle class home worldwide with blistering speed, many scholars predicted the emergence of a global consumer.

In the sixties, some scholars predicted that globalization and economic development would turn consumers in homogenous beings, resulting in the emergence of standardized marketing and advertising (Mooij, 2000). In particular, a group of scholars theorized that the decline in geographic and cultural differences would result in mass standardization (Roostal, 1963; Buzzell, 1968). One of the largest critics of cultural diversification was Harvard professor Ted Levitt. He posits that everyone’s wants and needs have homogenized and that as a result, marketing campaigns will also converge. Similarly, the world’s supply of goods and services will also homogenize (Levitt, 1983).

However, many studies have found that as economies and national wealth converge, i.e. less inequality, researchers continue to find that consumer culture becomes an increasingly important determinant of collective and individual needs and desires. (Hofstede, 2001; Mooij, 2000; Kacen & Lee, 2002; Singh, 2006)

**Objectives**

In this study, I will investigate the differences among consumers in a shopping environment, and whether cultural background affects their behavior. More specifically, I will differentiate between two shopping modes (browsers vs. goal directed shoppers) present in the shopping environment. Subsequently, I am interested in whether a
shopper’s cultural background influences which shopping mode they will be in. Furthermore, I investigate whether a shopper’s cultural background impacts purchase decision and time spent shopping given that they are in a goal directed shopping mode.

In short, I investigate the following:

1. Effects of shopping mode on
   a. Purchase incidence
   b. Time spent shopping
2. Effect of a consumer’s culture on which shopping mode they will likely be in
3. Interaction between culture and shopping mode on purchase incidence and time spent shopping.

Graphically, this looks like:

Contributions

This thesis will contribute in three distinct ways, namely (1) data superiority, (2) knowledge contribution, and (3) practical relevance. The data I will have access too is superior in the sense that it is unobtrusive, exceedingly accurate, unbiased, extensive, and complete. Previous data has often been collected through surveys, which despite its many advantages (e.g. measurability of unobservable constructs such as attitudes, perspectives, and perceptions) also suffers from known problems which might bias a study’s results. Other investigations have used voluntary RFID tags which also have significant limitations. As you will find out in the next section, the data I will have access to is accurate to 20 cm, it is unobtrusive and unbiased (people are not aware they are being observed), extensive in the sense that I have access to a large sample size, and complete because everyone with a mobile telephone is included.
My contributions will also be theoretical. To my knowledge, there is no other study which directly links uncertainty avoidance to a consumer’s propensity to browse or buy, nor on the existence of an interaction effect between goal directed shopping and uncertainty avoidance on purchase incidence.

Finally, my findings will be practically relevant for retailers and marketers. I will show that retailers should focus their efforts on specific customer segments. As a result, retailers should train their personnel to look for signs of goal directedness and browsing. Similarly, online retailers should tailor their web shop to take into account country of origin or location. Furthermore, as the world economy is becoming increasingly cross-cultural, it will become ever more important to understand what drives consumer behavior.

2. Conceptual Framework

In this section I will take a closer look at shopping modes, their impact on behavior, and Hofstede’s 5 dimensions of culture. I will also review empirical evidence to provide a strong link between my hypotheses and the body of literature.

Shopping Modes

I argue that a consumer will be in either of 2 shopping modes: goal directed shopping or browsing. Grounded on existing theories from marketing and consumer psychology, I define that the goal directed shopper is looking to make a specific purchase, will only visit the required store, get out as soon as possible, and is not interested in anything else than making the purchase he planned to. Contrarily, the browser enjoys the process of shopping, visits many different stores, spends more time per store and collects information for a future purchase. In the table below, I briefly summarize the findings of various papers before mine who have conceptualized shopping modes across different dimensions. The distinction between goal-directed shoppers and browsers, which I adopted, is most closely linked to the paper by Janiszewski (1998) who classifies a consumer as being either in a goal directed or exploratory search mode during a shopping trip.
<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Journal</th>
<th>Cites</th>
<th>(Most important) Typologies Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barry Babin, William Darden, Mitch Griffin (1994)</td>
<td>Journal of Consumer Research</td>
<td>1751</td>
<td><strong>Hedonic</strong>: subjective and personal, represents entertainment and/or emotional value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Utilitarian</strong>: being rational, task-related and often described as a chore or errand</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Apathetic shopper</strong>: has little or no attachment to shopping, views it as purely utilitarian or a chore</td>
</tr>
<tr>
<td>Jack Lesser, Marie Hughes (1986)</td>
<td>Business Horizons</td>
<td>59</td>
<td><strong>Active shopper</strong>: enjoy shopping, search for bargains</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Inactive shopper</strong>: has extremely limited shopping interests</td>
</tr>
<tr>
<td>Danny Bellenger, Pradeep Korgaonkar (1980)</td>
<td>Journal of Retailing</td>
<td>540</td>
<td><strong>Recreational shopper</strong>: spend more time per trip, visit different shops even after having made a purchase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Convenience shopper</strong>: knows in advance what they will purchase and rarely ever buy products they don’t immediately need</td>
</tr>
<tr>
<td>Danny Bellenger, Dan Robertson, Barnett Greenberg (1977)</td>
<td>Journal of Retailing</td>
<td>242</td>
<td><strong>Recreational shoppers</strong>: desire a high quality environment with a large number and variety of related products and services</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Convenience shoppers</strong>: want a shopping environment close to home with low prices and convenience</td>
</tr>
<tr>
<td>Michael Guiry, Anne Magi, Richard Lutz (2006)</td>
<td>Journal of the Academy of Marketing Science</td>
<td>60</td>
<td><strong>Shopping enthusiasts</strong>: visit the mall more frequently and spend significantly more time, also visit a higher amount of stores and spend more money</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Aversive shoppers</strong>: are the opposite from shopping enthusiasts in every dimension.</td>
</tr>
<tr>
<td>Chris Janiszewski (1998)</td>
<td>Journal of Consumer Research</td>
<td>211</td>
<td><strong>Exploratory search</strong>: (also known as browser) consumer is not necessarily focused on making a purchase but is more driven by hedonic considerations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Goal directed search</strong>: consumer has a specific purchase in mind and the consumer engages in a focused and directed pattern of search</td>
</tr>
</tbody>
</table>
Table 1 clarifies that there is no unified definition of different shopping modes. Prior scholars have focused on different constructs and taxonomies to refer to shoppers' goals during shopping trips. In this thesis, I propose a unified conceptualization of shopping modes which integrates findings from this rich literature.

Bellenger & Korgaonkar (1980) define the “recreational” shopper as spending more time per shopping trip and visiting different shops even after having made a purchase. “Convenience” shoppers on the other hand usually know in advance what they will purchase and rarely ever buy products they don’t immediately need (Bellenger & Korgaonkar, 1980). These conclusions mirror earlier findings by Bellenger et al. (1977), who theorize that “recreational” shoppers desire a high quality environment with a large number and variety of related products and services while “convenience” shoppers want a shopping environment close to home with low prices and convenience. Likewise, Bloch et al. (1994) find that shopping “enthusiasts” use a wide range of services the shopping environment has to offer, spend the most time and money in the mall, visit the highest number of stores and revisit the mall most often. Shopping “minimalists” participate least in all mall activities; they are uninterested in browsing, socializing, or consuming and view the mall as a hassle and just want to get their chores done. These conclusions are mirrored by Reynolds et al. (2002) and Guiry et al. (2006). By looking at their characteristics, we can see that recreational shoppers are very similar to the enthusiasts; likewise, convenience shoppers are similar to minimalists.

Finally, Janiszewski (1998) dichotomizes shopping behavior into goal directed and exploratory search. Exploratory search (also known as browser) refers to behavior where the consumer is not necessarily focused on making a purchase but is more driven by hedonic considerations. In other words, the consumer does not necessarily derive utility from a purchased product or service, but instead is seeking pleasure from the act of browsing itself (Bloch, Sherrell, & Ridgway, 1986). This pleasure derived from the browsing behavior is triggered by the consumer’s desire to attain knowledge related to products which may be useful in the future, and while purchasing is not the immediate goal, a purchase may still occur if the consumer is triggered by the right stimulus (Wendy W. Moe, 2001). As I have discussed earlier, these consumers
(referred to as enthusiasts, recreational and exploratory search) tend to spend a greater amount of time and money at stores. As Langrehr (1991) noted, purchases at a store are secondary to the shopping experience; “People buy so they can shop, not shop so they can buy.” Similarly, Bellenger & Korgaonkar (1980) find that shopping enjoyment and time spent are positively correlated due to the derived enjoyment of the process of shopping. Contrarily, goal directed search is triggered when a consumer has a specific purchase in mind and the consumer engages in a focused and directed pattern of search (Janiszewski, 1998). As we have seen earlier, these consumers (referred to as convenience, minimalist and goal directed) generally spend less time and money in-store. Shopping mode is not a stable trait. Shoppers oscillate or switch shopping mode in-between trips; while a shopper might be goal directed one day, he might be browsing the next (Sherry J., 1990).

I build upon Janiszewski’s findings and posit that browsers are consumers who enjoy the act of shopping, spend a significant amount of time and money doing so, and visit many different stores; this mode coincides with shopper typologies recreational shopper, enthusiasts, and exploratory search (Bellenger & Korgaonkar, 1980; Bloch, Ridgway, & Dawson, 1994; Guiry, Magi, & Lutz, 2006; Janiszewski, 1998). One of the primary goals of browsers is the gathering of product information which may become useful when making a purchase in the future.

Goal directed shoppers on the other hand see shopping as a mandatory chore to acquire a specific product, only visit the required store to make the purchase, and prefer to spend as little time as possible doing so; this mode coincides shopper typologies economic shopper and aversive shoppers (Bellenger & Korgaonkar, 1980; Guiry, Magi, & Lutz, 2006; Lumpkin, 1985; Reynolds, Ganesh, & Lucket, 2002; Westbrook & Black, 1985).

Concluding, goal directed shoppers have a specific product in mind, visit as few stores as possible and spend as little time as possible in order to fulfill their needs. Browsers however, enjoy the act of shopping, spend significantly more time and money, and visit more stores. While goal directed shoppers look to acquire specific products, browsers
are often interested in acquiring information which may be useful when making future purchases.

**Antecedents of a Consumer's Shopping Mode: The Role of Culture**

There are several determinants of shopping mode. The gender of a shopper is a significant determinant of shopping mode with females being much more likely to browse and shop for pleasure than males (Dholakia, 1999). Furthermore, socioeconomic class is a weak determinant of shopping mode; the higher up the social class the more frequently the consumer shops, however the less time they spend per trip and as a result, the middle class are most likely to browse (Rich & Jain, 1968). Finally, Jarboe et al. (1987) find that browsers also differ psychologically; they view themselves as opinion leaders, have greater confidence in themselves and view themselves as being high up the social class ladder.

Gender aside, demographics do not have the predictive power we are looking for in order to predict consumer shopping mode. There is a growing interest in finding new, alternate determinants of shopping mode. Culture impacts our lives in so many different ways from what kind of products we consumers to the way we behave at the dinner table. Although culture has not yet been explored as a predictor of in-store or in-mall shopping behavior, there is significant evidence to suggest it should be considered.

Grounded on existing theories in marketing and cultural psychology I will postulate that a consumer's cultural background plays an important role in their behavior in the shopping environment.

**Defining Culture**

Culture has been defined in many different ways. A commonly accepted definition is that by Linton (1945), which affirms culture as a set of learned behaviors common to members of a particular society. One of the most widely accepted frameworks of culture has been put forward by Geert Hofstede (2001) in the 1980s and have been revised many times and replicated over 1600 times since.

Hofstede's (2001) dimensions of culture are derived from data collected at IBM, from employees all around the world; the results come from extensive data, namely 116,000
questionnaires from over 60,000 respondents from 70 different countries. His work is widely recognized and his empirical results have been imitated by many others (Kacen & Lee, 2002; Michon & Chebat, 2004; Mooij, 2000; Shim & Gehrt, 1996; Singh, 2006).

Based on these findings, Geert Hofstede operationalized culture in the form of 5 dimensions, including power distance (PDI), individualism vs. collectivism (IDV), masculinity vs. femininity (MAS), uncertainty avoidance (UAI), and long term vs. short term orientation (LTO). Furthermore, he indexed all participating nations and linked these indexes with economic, demographic, political, and geographic aspects of society (Soares, Farhangmehr, & Shoham, 2007). As a result, it is the most robust and comprehensive framework which has become the norm for international marketing practitioners performing cross cultural studies (Samiee & Jeong, 1994).

First I will review Hofstede’s 5 dimensions of culture, after which I will mention the relevance to my research.

Individualism – collectivism represents the degree to which either the individual, or the group as a collective is important in society. In collectivist culture, the group is important, and large extended families are an observable attribute of collective cultures. In individualistic cultures, the individual and his or her most direct family is most important. The most individualistic culture is that of the United States, with a score of 91, with Venezuela being the least individualistic, scoring 12.

Power distance reflects the degree to which people in society accept hierarchy, authority, and power inequality. High power distance countries (e.g. Malaysia) accept that people with a higher formal ranking either at work or anywhere else have more power than their subordinates. Low power distance countries (E.g. United Kingdom) people regard each other more as equals regardless of rank of status (Hofstede, 2001).

Masculinity refers to the degree to which countries adopt either more masculine values, or less masculine values (referred to as feminine). High masculine countries such as Japan place higher emphasis on such as competitiveness, ambition, power, and materialism. Feminine countries such as Sweden regard life quality, relationships, and social goals (Hofstede, 2001).
Long-versus short-term orientation formerly referred to as the “Confucian Dynamism,” defines refers to the importance of time horizon within a society. More specifically, societies with a long term orientation (e.g. China) place emphasis on the future, emphasizing adaptation, saving, and persistence. Contrarily, short term oriented societies (often western societies) have a desire to establish truth, wish to champion short term success and have deep respect for traditions (Hofstede, 2001).

Uncertainty avoidance relates to the degree to which a culture is uncomfortable or feels threatened by uncertainty or doubt. High uncertainty avoidant cultures prefer to avoid as much uncertainty as possible. The most uncertainty avoiding culture is that of Greece, scoring 112, while Denmark scores the lowest at 23. The United Kingdom scores 35, Germany 65, France 86, Spain 86 and Portugal 104 on the uncertainty avoidance scale (Hofstede, 2001).

**Culture as a Predictor**

I wish to investigate whether people from uncertainty avoiding countries are more likely to be in the information seeking stage than people from more uncertainty accepting cultures.

Consumer behavior is affected by consumer uncertainty (Urbany, Dickson, & Wilkie, 1989), and consumer uncertainty motivates information sourcing behavior. Some have found product uncertainty, knowledge, and/or choice uncertainty to be determinants of purchase probability (Dash, Schiffman, & Berenson, 1976; Gunasti & Ross, 2009); what about a consumer’s uncertainty avoiding nature?

On average, consumers prefer certainty over uncertainty (Urbany, Dickson, & Wilkie, 1989), and uncertainty leads to higher levels of search (Lee & Qiu, 2011). In an investigation into the effects of an informational website on in-store revenues, Pauwels et al. (2011) find that for a specific segment, “smart fans,” it has a positive significant effect. Smart fans are highly responsive to price and product fit; this segment is concerned with making the right choice and therefore engage in information seeking behavior before making an actual purchase (Pauwels, Leeflang, Teerling, & Huizingh, 2011). This notion is consistent with our idea of an uncertainty avoiding consumer, who
prefers to research before making a product acquisition in order to reduce uncertainty regarding alternative products and prices. Accepting the fact that certain countries prefer to avoid uncertainties to a higher extent than other countries, I posit that uncertainty avoidance acts as a predictor for shopping mode.

As a result, I hypothesize the following:

\[ H_1: \text{Consumers from countries with high uncertainty avoidance will more likely be in a browsing shopping mode than in a goal-directed shopping mode compared to consumers from countries with low uncertainty avoidance.} \]

Consequences of Consumer’s Shopping Mode: Purchase Incidence and Time Spent in Store

I will now theorize about the effects of shopping mode (browsing vs. goal directed) on shoppers’ (a) purchase likelihood, and (b) time spent in store.

Effects of Shopping Mode on Purchase Incidence

Previously, I explored different shopping modes and differentiated between the “browser” and the “goal directed shopper.” I also showed you that shopping mode is not a stable trait as it can fluctuate over different shopping trips. I am interested in whether shopping mode (browsing vs. goal directed) provides any indication of a shoppers purchase incidence. More specifically, I will argue that browsers are more likely to purchase than goal directed shoppers.

We have seen that browsers spend a higher amount of money during their shopping trip than do goal directed shoppers (Bloch, Ridgway, & Dawson, 1994; Guiry, Magi, & Lutz, 2006; Reynolds, Ganesh, & Lucket, 2002). Research done specifically to investigate the relationship between sales volume and shopping path length provides even more support that sales volumes are positively and significantly correlated with shopping path length (Kholod, Nakahara, Azuma, & Yada, 2010).

On the one hand, I would expect that the goal directed shopper has a higher purchase incidence because based on his or her motivations it would seem that he or she has already made up their mind on buying that bottle of water. On the other hand, I expect
that since the browser is passing more shelf space and looking at more products, the browser might be more in the ‘mood’ to make purchases. Furthermore, we have already seen that browsers spend more money.

Shim et al. (2001) propose that the intention to search for product information in an online shopping environment is an important indicator of purchase intention. They suggest that search intention plays a mediating role between purchase intention and various other important antecedents of purchase intention such as perceived and actual behavioral control as specified by Ajzen’s (1991) Theory of Planned Behavior. Klein (1998) posits in her “Interaction Model of Pre-purchase Consumer Information Search” that consumers only cease to search for information when the costs of search outweigh the benefits. Similarly, in order to complete compound shopping tasks, consumers develop shopping strategies which contain any number of steps (Darden & Dorsch, 1990). In other words, a specific shopping goal is completed by walking through a number of connected steps or acts. This proposition is alike to Gollwitzer’s (1993) concepts which state that such explicit strategies lead to a decision to either complete or not complete the desired behavior. Orbell et al. (1997) state that consumers who go through this process are more likely to fulfill these intentions (to purchase) than those who do not. We have seen that browsers are explicitly acquiring information for possible upcoming purchase decisions; this could be in 5 minutes or 5 days’ time. As a result, I posit that browsers will have a higher purchase incidence than goal directed shoppers.

Empirically Hui et al. (2009) find that purchase behavior is interconnected with patterns of visitation. More specifically, consumers who visit a higher amount of different sections within the supermarket have a higher purchase incidence (Hui, Bradlow, & Fader, 2009). In an online environment, Shim et al. (2001) find a significant and positive effect between information search and purchase intention. This reinforces my postulation that browsers have a higher probability of making a purchase than do goal directed shoppers.

Based on the literature, I hypothesize that:
$H_{2A}$: Consumers in a browsing shopping mode, will exhibit the highest purchase incidence than consumers in a goal directed shopping mode.

**Effects of Shopping Mode on Time Spent Shopping**

We have established that browsers are likely to visit more stores, spend more money, and most importantly, spend more time shopping; their goal is to collect product information in order to (possibly) make a future purchase. This is because browsers generally enjoy shopping, view it as a fun pastime, and receive gratification from the act of shopping (Bloch, Ridgway, & Dawson, 1994; Guiry, Magi, & Lutz, 2006; Reynolds, Ganesh, & Lucket, 2002).

Hirschman & Holbrook (1982) posit that consumers engaged in fun activities exhibit a higher need for emotional and cognitive attachment and therefore expend a larger amount of time doing so. Similarly, others recognize the relationship between emotional state, time spent shopping, and purchase incidence (Dawson, Bloch, & Ridgway, 1990; Spangenburg, Crowley, & Henderson, 1996; Kellaris & Kent, 1993). Subsequently, the Mehrabian – Russell model states that emotional states are triggered by feelings of pleasure, arousal, and dominance (Mehrabian & Russell, 1974). As a result, we can expect that shoppers who extract pleasurable feelings from the act of shopping are more likely to be in an emotional state which leads them to spend more time fulfilling the act. Empirically, Bellenger & Korgaonkar (1980) find that shopping enjoyment and time spent go hand in hand. Similarly, Schmidt & Spreng (1996) theorize that shopping enthusiasm (enjoyment derived from shopping) provides a strong motivation to seek information for the intrinsic delight of the shopping process itself. Others have found that factors such as in-store background music alter a consumers’ state of mind and actually induces them to shop longer (Yalch & Spangenberg, 2000).

Other academics have also found an inverse relationship between knowledge and search (Anderson, Engledow, & Becker, 1979; Moore & Lehmann, 1980; Newman & Staelin, 1971; Swan, 1969). One explanation might be that goal directed shoppers have acquired all the information they need and therefore do not engage in (as much) search activities as browsers (Brucks, 1985). This would explain the positive relationship between time spent shopping and knowledge uncertainty Urbany et al.
knowledge uncertainty is defined by them as all dimensions of knowledge a consumer needs in order to make a purchase.

In short, browsers will spend more time shopping per store due to their enjoyment of the experience. However, their drive to become more knowledgeable in order to make a more well-informed purchase in the future is also a key factor. Based on these findings, I hypothesize that:

\[ H_{2B} \]: Consumers in a browsing shopping mode will spend (on average) the most amount of time shopping per store compared to consumers in a goal directed shopping mode.

The Moderating Effect of Culture on the Relationship between Shopping Mode and Shopper Behavior

Previously, I argue that high uncertainty avoiding consumers are more likely operate in a browsing mode. This does not imply that high uncertainty avoiding cultures are less likely to purchase; after all, I previously argued that browsers have a higher purchase incidence than goal directed shoppers. I will investigate whether uncertainty avoidance moderates the effect of shopping mode on shopper behavior, namely on its effect on (i) purchase incidence and (ii) time spent in store.

Moderating Effect of Culture on the Relationship between Shopping Mode and Purchase Incidence

I first focus on the moderating effect of culture on the relationship between shopping mode and purchase incidence. In other words, given that a consumer is in a goal directed shopping mode, I want to study whether she is more or less likely to make a purchase if she comes from a country whose culture is characterized by higher levels of uncertainty avoidance.

Culture has been a proven antecedent to many different forms of consumer behavior; from purchase habits, to preference for new over second hand cars. But how does culture act as a moderating factor? Hewett et al. (2002) posit culture moderates the effect between buyer perception and purchase intention. Similarly, uncertainty
avoidance has been known to affect the purchase of insurance, second-hand products, and beauty products all in order to mitigate risk (Marieke de Mooij, 2002).

Being familiar with a product (i.e. having acquired enough information regarding it) has significant negative effects on decision time to purchase that product; high familiarity leads to a quicker purchase (Park & Lessig, 1981). Similarly, Laroche et al. (1996) posit that a consumer’s intention to purchase is significantly and positively related to their attitude or confidence. Attitude and confidence are affected by the consumer’s familiarity with the product, findings mirrored by Laroche & Sadokierski (1994). Thus, as a consumer becomes familiar with a product, they are more likely to purchase.

We have seen that consumer behavior is not a stable trait. When consumers gain experience from browsing and purchasing, their perceptions and behaviors change accordingly (Yu, Ha, & Rho, 2005). I have also reviewed literature which demonstrates that consumers follow a set of steps in order to complete shopping tasks (Darden & Dorsch, 1990). Such shopping strategies lead to either a purchase or not (Gollwitzer, 1993). A critical factor in the purchase process is trust between the consumer and the retailer. In particular, in situations of uncertainty, uncertainty is reduced by trust and apparent risk is decreased (Doney, Cannon, & Mullen, 1998). Trust is defined by Moorman et al. (1993) as "a willingness to rely on an exchange partner in whom one has confidence." A culture’s uncertainty avoidance is the degree to which a culture is uncomfortable or feels threatened by uncertainty or doubt; thus we would expect high uncertainty avoiding culture’s to be less willing to make purchase a certain product until he is absolutely sure he knows what he wants. I have defined goal directed consumers as those consumers out to make a specific purchase; they already have the information they need. Thus, high uncertainty avoiding customers who are in a goal directed shopping mode have gained enough insights in order to adequately make a purchase decision. I posit that these consumers will have a higher purchase incidence than uncertainty accepting customers, because of their readiness to engage in a goal directed shopping mode signals their intent that they have collected enough information and reduced uncertainty to acceptable levels. As a result, they are ready to make a purchase.
Thus, I hypothesize that:

$H_{3A}$: **Consumers from countries with high uncertainty avoidance who are in a goal-directed shopping mode will have a higher purchase incidence relative to consumers from countries with low uncertainty avoidance (in a goal directed shopping mode).**

**Moderating Effect of Culture on the Relationship between Shopping Mode and Time Spent Shopping**

As shown by Park and Lessig (1981), consumers familiar with a product will spend less time shopping. I also discussed that shopping is an uncertain process and that consumers from high uncertainty avoiding cultures are less likely to engage in uncertain action than those from uncertainty accepting cultures. As a result, I postulate that consumers from uncertainty avoiding cultures who engaged in goal directed shopping have a stronger intention to make a purchase than consumers from countries with low uncertainty avoidance cultures. Thus, I theorize that these consumers will need less time making the actual purchase than uncertainty accepting cultures.

$H_{3B}$: **Consumers from countries with high uncertainty avoidance who are in a goal-directed shopping mode will (on average) spend less time shopping per store relative to consumers from countries with low uncertainty avoidance.**
Overview of Hypotheses

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H₁</strong></td>
<td>Consumers from high uncertainty avoiding countries will more likely be browsing than in a goal directed shopping mode compared to people from uncertainty accepting country.</td>
</tr>
<tr>
<td><strong>H₂A</strong></td>
<td>Consumers in a browsing shopping mode will exhibit the highest purchase incidence than consumers in a goal directed shopping mode.</td>
</tr>
<tr>
<td><strong>H₂B</strong></td>
<td>Consumers in a browsing shopping mode will spend (on average) the most amount of time shopping per store compared to consumers in a goal directed shopping mode.</td>
</tr>
<tr>
<td><strong>H₃A</strong></td>
<td>Consumers from countries with high uncertainty avoidance who are in a goal-directed shopping mode will have a higher purchase incidence relative to consumers from countries with low uncertainty avoidance.</td>
</tr>
<tr>
<td><strong>H₃B</strong></td>
<td>Consumers from countries with high uncertainty avoidance who are in a goal-directed shopping mode will (on average) spend less time shopping per store relative to consumers from countries with low uncertainty avoidance.</td>
</tr>
</tbody>
</table>

3. Methodology

In this section we will take a look at the data, what has been done to the data, where it came from, and how we will use the data to find evidence for or against the hypotheses.

Data

I estimate my model from data collected through a shopping path tracking technology developed by a new tech start-up and recorded at a medium-sized European airport (the airport of Porto in Portugal, used by about 6 million passengers in 2011). The firm uses a set of antennae to track mobile devices in a physical area. The antennae are capable of tracking GSM (2G and 3G), Wi-Fi, and Bluetooth signals of devices such as cell phones and tablets. Individuals only have to enter the area with a functional mobile telephone to be tracked. Consumer movement is tracked by recording a set of coordinates every 5 seconds, recorded in a database.

The antennae record the device’s MAC address which is unique to one particular device, as ID. Thus, if an individual returns 2 times in a week, they are recognized on
their second visit. Interestingly enough, there are no repeat visitors. Individuals are tracked once entering the security zone until they have left the duty free shopping zone as seen below.

A, B, C, and D represent individual stores, and the X’s represent the individual antennae. Green areas are frequently visited areas by individuals; the red areas indicate the most popular areas.
<table>
<thead>
<tr>
<th>Store</th>
<th>Hours</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>07.00 – 23.00</td>
<td>Regional products (jams, honey, olive oil, ham, wines, biscuits, ham &amp; other regional products...)</td>
</tr>
<tr>
<td>B</td>
<td>07.00 – 23.00</td>
<td>Fashion &amp; Clothing for male and female</td>
</tr>
<tr>
<td>C</td>
<td>07.00 – 23.00</td>
<td>Fashion, consumer electronics and branded accessories</td>
</tr>
<tr>
<td>D</td>
<td>07.00 – 23.00</td>
<td>Gifts and products such as chocolates, tobacco and liqueur and hedonic gifts such as cosmetics and perfumes</td>
</tr>
</tbody>
</table>

Between January 15, 2012 and January 28, 2012 a total of 45,980 mobile devices traversed through the observed area. The mobile devices of our sample originated from only 5 different countries, namely Portugal (27,000), Spain (6,767), France (3947), the UK (3,543), and Germany (3,469).
Respondents | 45,980  
---|---
- Excluding non-shoppers | 23,287  

| Average # of Stores Visited | 0.55 Stores  
---|---
- Excluding non-shoppers | 1.35 Stores  

| Average time spent per store | 151 seconds (02:31)  
---|---
- Excluding non-shoppers | 304 seconds (05:04)  
- Browsers | 356 seconds (2599 consumers)  
- Goal Directed | 297 seconds (19749 consumers)  

| Nationalities (Origin of cell phone) | 5 (Portugal, Spain, France, UK, and Germany)  
---|---

| Store A Visits | 4847  

| Store B Visits | 4520  

| Store C Visits | 5029  

| Store D Visits (Wine & Others) | 13541  

| Store D Visits (Cosmetics) | 10913  

| Store A Purchases | 1105  

| Store B Purchases | 927  

| Store C Purchases | 1230  

| Store D Purchases (Together) | 5993  

---

1 I define shoppers as people who visited at least 1 store
Variable Operationalization

Before discussing the actual model, I will operationalize my variables.

<table>
<thead>
<tr>
<th>Variable (Abbreviation)</th>
<th>Description/Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Spent Shopping ($TSP$)</td>
<td>TSP is represented in seconds, and displays the amount of time a consumer spends on average per store. It is calculated by dividing the total time spent shopping by the amount of stores visited.</td>
</tr>
<tr>
<td>Purchase Incidence ($PI$)</td>
<td>$PI$ is a dummy variable whether a consumer has made a purchase or not. It takes the value 1 or 0 respectively.</td>
</tr>
<tr>
<td>Browsing ($Br$)</td>
<td>$Br$ will be operationalized by a dummy variable, indicating whether a consumer visits all four stores (browsing), or just one (goal directed). So $Br$ will take on the value 1 for browsers (those who visited four stores), and the value 0 for goal-directed shoppers (those visiting 1 store).</td>
</tr>
<tr>
<td>Uncertainty Avoidance ($UA$)</td>
<td>$UA$ will take the form of 4 dummy variables. The base case will represent consumers originating from the UK which have the lowest uncertainty avoidance index. Subsequently we will have 4 dummy variables representing consumers originating from Germany, France, Spain and Portugal. Uncertainty avoidance will also be specified as high vs. low uncertainty avoidance. The United Kingdom scores 35, Germany 65, France 86, Spain 86 and Portugal 104; Portugal will be our base case (no dummy), Spain and France will be classified as Medium, and the UK and Germany as Low.</td>
</tr>
</tbody>
</table>

Econometric Model

To test my hypotheses, I will conduct a linear regression model, a logistic regression model, and a chi-square test. I used the PASW statistical software to estimate my models. My dependent variables will be time spent shopping per store (linear regression), purchase incidence (logistic regression), and shopping mode (chi-square model).
The independent variables will include uncertainty avoidance and shopping mode.

For H1, I will compare the percentage of consumers in a particular shopping mode across a number of different countries. To realize this, I will use a chi-squared test which generalizes a two-sample Z-test in order to handle more than two parts. To complement this, I will use a logistic regression with shopping mode as the dependent variable, and dummy variables for the country of origin as the independent variables.

For hypothesis H2A and H3A I will use a logistic regression model with purchase incidence (PI) as the dependent variable and uncertainty avoidance (UA), browsing (Br), and an interaction effect between uncertainty avoidance and shopping mode as independent variables. The model (model II) will look like this:

\[ PI_i = \begin{cases} 1 & \text{if the } i^{th} \text{ shopper made a purchase} \\ 0 & \text{otherwise} \end{cases} \]

In line with standard logistic regression, I assume that the fact that the shopper decides to make a purchase \((PI_i=1)\), indicates that the latent utility, for the consumer, of the purchase exceeds the latent utility of no purchase. For simplicity I denote the difference between the latent utility of purchase vs. no purchase as \(PI_i^*\) and define it as:

\[ PI_i^* = \alpha + \beta_{1,1} * UA_i + \beta_{2,1} * Br_i + \beta_{3,1} * UA_i * Br_i + \epsilon_i \]

For hypothesis H2B and H3B I will use a linear regression model with time spent per store (TSP) as dependent variable, and uncertainty avoidance (UA), browsing (Br) and an interaction between uncertainty avoidance and shopping mode as independent variables. The model (model III) will look like this:

\[ TSP_i = \alpha_2 + \beta_{1,2} * UA_i + \beta_{2,2} * Br_i + \beta_{3,2} * UA_i * Br_i + \epsilon_{i2} \]

**Data Cleaning**

We start with 45,980 cases, and after removing everyone who did not spend at least 1 second in any given store, we are left with 22,358 cases which seem to be clustered around 2 distinct groups. We create a number of variables for our regressions; the most
important being time spent per store. After analyzing boxplots and using Mahalanobis $D^2$ to investigate univariate and multivariate outliers respectively, we remove another 9 cases and are left with 22,348 cases for our final analysis.

For the complete process of data cleaning, I refer you to the section in the appendix marked data cleaning.

4. Results

Testing Assumptions

In our analysis, we will use logistic regression, linear regression, and a chi-squared test in order to test our hypotheses. Multicollinearity however can provide implausible results (Field, 2005). For linear regression to be BLUE, we need to fulfill more stringent assumptions such as linearity, normality, homoscedasticity, and independence (Field, 2005). The assumptions for chi-squared are lighter in nature: the sample must be selected from the population at random, and the sample size needs to be large enough so that the minimum number of observations in an individual category is at least five (Field, 2005).

We start with the assumptions logistic regression. We test for multicollinearity using a correlation matrix we find that the only two variables that are highly correlated (>0.8) are number of visits and time spent per store; however these variables never appear as independent variables in the same model. I also observe the standard errors during the analysis, and remark that all S.E. are below two, indicating no numerical errors such as multicollinearity (Field, 2005).

For linear regression (model 2) there are some stricter assumptions which need to be met. These assumptions include: linearity, homoscedasticity, normality and independence. We will test each of the assumptions in the same order.

Linearity and homoscedasticity can be tested by analyzing a residual plot with predicted values on the x axis and standardized residuals on the y axis. Linear data can be
observed by the symmetrical distribution around a horizontal line. Heteroskedasticity can be spotted by a cone shape fanning out to the right. Looking at our scatterplot, we see that there is no non-linear relationship present; however the presence of heteroskedasticity is ambiguous as the plot seems to suggest that the error variance is larger for lower values of the standardized predicted value, raising suspicion regarding heteroskedasticity. In order to better understand this and formally test for the presence of heteroskedasticity, I will run a Breusch-Pagan and Koenker test (Griffiths, Hill, & Lim, 2008). SPSS does not support these test in their GUI, however using a syntax file developed by Marta Garcia-Granero and available from her website, I run these tests using text commands. Looking at the results, we see significant evidence of the presence of heteroskedasticity. As a result, I will use a confidence interval of 1% instead of 5% and use a log transformation where possible (Breen, Jagannathan, & Ofer, 1986). The log transformations improve the situation slightly but heteroskedasticity still persists. Using a smaller confidence interval however, is a form of standard error correction since weighted least squares and White’s correction are not available in PASW (Breen, Jagannathan, & Ofer, 1986).

In order to test the assumption of normality, we look at the distribution of the error terms using a histogram. Observing our histogram of residual distribution, we see that the data is not completely normally distributed; however, it is not a bad fit either. Ideally, we might try a non-linear transformation, but considering our predictor variables are 1’s and 0’s, this will not be very effective.

**Results**

In this section I will run each model one by one and discuss the results. I have arranged all of the data in tables inserted in the text. To see the original PASW output, see appendix. I will discuss the findings per model; first I will discuss model I \((H_1)\), followed by model II \((H_{2A} & H_{3A})\) and finally model III \((H_{2B} & H_{3B})\). You will see that we find strong support for \(H_{2A}, H_{2B}\), some support for \(H_1\) and no support for \(H_{3A}\) and \(H_{3B}\).

**Model I**

The chi-squared test results indicate (below) that there are some differences in shopping mode among the different countries of origin; the Germans, Spaniards, and
Portuguese browse more than we would expect while the French and English browse less. To test the significance of these findings we look at the (two-tailed) Pearson Chi-Square significance which at 0.096 is larger than 0.05 (but smaller than 0.10) indicating that there is still some probability that these findings are down to chance. Thus we cannot yet significantly conclude that there is a difference between nations regarding their propensity to either browse or be goal directed, as the difference we observe could attributed to chance.

<table>
<thead>
<tr>
<th>Goal Directed</th>
<th>Count Expected Count</th>
<th>Count</th>
<th>Spain</th>
<th>France</th>
<th>Portugal</th>
<th>U.K.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>1517</td>
<td>2940</td>
<td>1688</td>
<td>11963</td>
<td>1641</td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>Spain</td>
<td>1520.9</td>
<td>2949.8</td>
<td>1685.2</td>
<td>11988.3</td>
<td>1604.8</td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>France</td>
<td>1588</td>
<td>1603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>Portugal</td>
<td>11963</td>
<td>1641</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>U.K.</td>
<td>1603</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>Total</td>
<td>19749</td>
<td>19749</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19749</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Browsing</th>
<th>Count Expected Count</th>
<th>Count</th>
<th>Spain</th>
<th>France</th>
<th>Portugal</th>
<th>U.K.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>204</td>
<td>398</td>
<td>219</td>
<td>1603</td>
<td>175</td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>Spain</td>
<td>200.1</td>
<td>388.2</td>
<td>221.8</td>
<td>1577.7</td>
<td>211.2</td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>France</td>
<td>1603</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>Portugal</td>
<td>1603</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>U.K.</td>
<td>175</td>
<td>211.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>Total</td>
<td>2599</td>
<td>2599</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2599</td>
</tr>
</tbody>
</table>

Before conducting a binary logit model, I try a different specification of culture. Previously I looked at individual countries as predictors, now I will compare low (UK and Germany) vs. medium (France and Spain) vs. high (Portugal) uncertainty avoiding countries. Running the chi-square test again yields the following results.

<table>
<thead>
<tr>
<th>Goal Directed</th>
<th>Count Expected Count</th>
<th>Count</th>
<th>Low</th>
<th>Med</th>
<th>High</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>3158</td>
<td>4628</td>
<td>11963</td>
<td></td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>Spain</td>
<td>3125.7</td>
<td>4635</td>
<td>11988.3</td>
<td></td>
<td></td>
<td>19749</td>
</tr>
<tr>
<td>France</td>
<td>379</td>
<td>617</td>
<td>1603</td>
<td></td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>Portugal</td>
<td>411.3</td>
<td>610</td>
<td>1577.7</td>
<td></td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>U.K.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2599</td>
</tr>
<tr>
<td>Total</td>
<td>19749</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19749</td>
</tr>
</tbody>
</table>

To test the significance of these findings we look at the (two-tailed) Pearson Chi-Square significance which at 0.180 is larger than 0.05, and almost twice as large as the previous specification, indicating that it is very probable that these findings are down to
chance. Thus we cannot yet significantly conclude that there is a difference between cultures, as the difference we observe could be chance.

To investigate this further, we will use a binary logit model to complement the chi-squared test. The chi-squared statistic of the Hosmer-Lemeshow test is close to 0, indicating that the null-hypothesis of a good fit cannot be rejected. The presence of a relationship between dependent and independent variables is confirmed by the omnibus test of model coefficients. Looking at the results of the binomial logit model (below), we see that the fact that a consumer reside in Spain, Germany, and Portugal is a significant predictor of shopping mode, while being from France is significant only at 10% confidence interval (UK is our base case). Looking at Exp(B), we conclude that consumers from the UK are roughly 26% more likely to be in a browsing stage than Germans and Portuguese, and even more than the Spanish consumers. Consumers from the UK are roughly 21% more likely to browse than the French at 10% confidence interval.

<table>
<thead>
<tr>
<th>Dependent variable: dum_browse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>constant</td>
</tr>
<tr>
<td>France*</td>
</tr>
<tr>
<td>Germany**</td>
</tr>
<tr>
<td>Portugal**</td>
</tr>
<tr>
<td>Spain**</td>
</tr>
</tbody>
</table>

*indicates p<0.10  
**indicates p<0.05

Similarly to the chi-square test, I continue to run the model with the new uncertainty avoidance specification. We observe a good fit of the model to the data as indicated by the Hosmer-Lemeshow test, and the presence of a relationship between dependent and independent variables is confirmed by the omnibus test of model coefficients. I get the following results.
**Dependent variable: dum_browse**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant**</td>
<td>2.012</td>
<td>0.043</td>
<td>7.501</td>
</tr>
<tr>
<td>Low UA*</td>
<td>0.110</td>
<td>0.061</td>
<td>1.117</td>
</tr>
<tr>
<td>Med UA</td>
<td>-0.005</td>
<td>0.050</td>
<td>0.995</td>
</tr>
</tbody>
</table>

*indicates p<0.10  
**indicates p<0.05

We see that now uncertainty avoidance is characterized by high, medium, or low, the effect of origin practically disappears.

Hypothesis 1 questions whether consumers from high uncertainty avoiding countries are more likely be browsing than in a goal directed shopping mode compared to people from low uncertainty avoiding countries. The United Kingdom scores 35, Germany 65, France 86, Spain 86 and Portugal 104 on the uncertainty avoidance scale (Hofstede, 2001). While I find (moderate to weak) evidence to support that country of origin has an effect on a person’s shopping mode, the relation is different than expected. Where I predicted that consumers from countries characterized by low uncertainty avoidance would be more likely to be in a goal directed shopping mode, I actually found the opposite. The UK is the least uncertainty avoiding country, and also the most likely to browse. Similarly, Portugal and Spain are the most uncertainty avoiding countries, but have the lowest propensity to browse. When the model was specified differently (i.e. Low vs. High uncertainty avoidance instead of country of origin) we found that the effect disappears. I discuss possible reasons for not finding uncertainty avoidance as a predictor after discussing all three models.

**Model II**

Using the Hosmer-Lemeshow goodness of fit test in order to assess the models fit to the data, I observe that the chi-squared statistic is close to 0, indicating that the null-hypothesis of a good fit cannot be rejected. Furthermore, the presence of a relationship
between dependent and independent variables is confirmed by the omnibus test of model coefficients. Hence, we continue to the table of results.

**Dependent variable= purchase incidence (1/0)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.660</td>
<td>0.581</td>
<td>0.517</td>
</tr>
<tr>
<td>France</td>
<td>0.006</td>
<td>0.070</td>
<td>1.006</td>
</tr>
<tr>
<td>Spain</td>
<td>0.044</td>
<td>0.062</td>
<td>1.045</td>
</tr>
<tr>
<td>Portuguese</td>
<td>0.057</td>
<td>0.053</td>
<td>1.059</td>
</tr>
<tr>
<td>German</td>
<td>0.081</td>
<td>0.072</td>
<td>1.084</td>
</tr>
<tr>
<td>Browsing**</td>
<td><strong>0.832</strong></td>
<td><strong>0.181</strong></td>
<td><strong>2.299</strong></td>
</tr>
<tr>
<td>French*Browser</td>
<td>-0.190</td>
<td>0.240</td>
<td>0.827</td>
</tr>
<tr>
<td>Spanish*Browser</td>
<td>-0.235</td>
<td>0.216</td>
<td>0.791</td>
</tr>
<tr>
<td>Portuguese*Browser</td>
<td>-0.148</td>
<td>0.191</td>
<td>0.862</td>
</tr>
<tr>
<td>German*Browser</td>
<td>-0.047</td>
<td>0.249</td>
<td>0.954</td>
</tr>
</tbody>
</table>

**indicates significance at 5% CI**

In our model we have 2 main effects: country of origin (country dummies), and browsing (vs. goal-directed shopping mode), each either a 1 or a 0. We also have interaction effects, between browsing and country of origin. What we see is that browsing is the only significant independent variable, Browsing=1 indicates that a consumer is in a browsing mode meaning he has visited all 4 stores. The Exp(B) coefficient of 2.299 indicates that a 1 unit increase in Browsing increases the odds of making a purchase by 2.299 times. In other words, a browser is 229% more likely to make a purchase than a goal directed shopper. Furthermore, we find no significant main- or interaction-effects for uncertainty avoidance. Rerunning the model with uncertainty avoidance characterized as high vs. low, we find the following.

**Dependent variable= purchase incidence**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>-1.207</td>
<td>0.159</td>
<td>0.299</td>
</tr>
<tr>
<td>Browsing**</td>
<td><strong>0.882</strong></td>
<td><strong>0.167</strong></td>
<td><strong>2.417</strong></td>
</tr>
<tr>
<td></td>
<td>Low UA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>0.206</td>
<td>0.148</td>
<td>1.229</td>
</tr>
<tr>
<td>Low UA * Browsing</td>
<td>-0.198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High UA * Browsing</td>
<td>-0.071</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**indicates significance at 5% CI

Browsing has a significant effect on purchase effect while uncertainty avoidance has no effect. Browsers are now 2.417 times as likely to make a purchase compared to goal directed shoppers. We can see that the results change only (very) slightly when we rerun the model with a different specification of culture, attesting to the robustness of my findings.

The final robustness check I conduct is to test the effect of expatriates in Portugal (see next section). In order to minimize the effect of these expats, I exclude all Portuguese phones, and rerun both models. In both cases, the results stay the same.

In literature, many different academics theorize and find (empirically) that browsers spend more money and more time shopping than goal directed shoppers (Bloch, Ridgway, & Dawson, 1994; Guiry, Magi, & Lutz, 2006; Reynolds, Ganesh, & Lucket, 2002). This is attributed to the shopper’s intrinsic interest in shopping as an act alone, and because browsing for information is a clear signal of purchase intention (Shim, Eastlick, Lotz, & Warrington, 2001). My findings confirm those of many academics before me; I find strong support for H2A: browsers exhibit higher purchase incidence than goal directed shoppers.

We do not however, find any evidence for hypothesis H3A, i.e. purchase incidence does not seem to be affected by an interaction effect of uncertainty avoidance and browsing behavior. The reasons for this are similar to those for model I and will be discussed after the next model.

**Model III**
Looking at the results of our regression, I observe extremely high standard errors. As a result, I will take the natural logarithm of time spent shopping in order to alleviate this
problem. The $R^2$ of the regression is 0.142 indicating that the model explains 14.2% of the variance in the time spent shopping, which being only a moderate fit to the data, can be considered impressive given the parsimony of my model (basically explaining time spent shopping only by (i) shopping style and (ii) culture). We see the results below:

Dependent variable = LN(Time Spent Shopping)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>4.638</td>
<td>0.019</td>
</tr>
<tr>
<td>Browsing</td>
<td>2.576</td>
<td>0.054</td>
</tr>
<tr>
<td>France</td>
<td>-0.045</td>
<td>0.053</td>
</tr>
<tr>
<td>Spain</td>
<td>-0.035</td>
<td>0.042</td>
</tr>
<tr>
<td>U.K.</td>
<td>-0.054</td>
<td>0.054</td>
</tr>
<tr>
<td>Germany</td>
<td>-0.004</td>
<td>0.056</td>
</tr>
<tr>
<td>France*Browse</td>
<td>-0.013</td>
<td>0.156</td>
</tr>
<tr>
<td>Spain*Browse</td>
<td>0.065</td>
<td>0.122</td>
</tr>
<tr>
<td>U.K.*Browse</td>
<td>0.043</td>
<td>0.171</td>
</tr>
<tr>
<td>Germany*Browse</td>
<td>0.036</td>
<td>0.162</td>
</tr>
</tbody>
</table>

*** indicates significance at 1% CI

We see that only browsing is significant (at 1%), while there is no main effect of origin, nor an interaction effect between origin and browsing. The dependent value is the natural logarithm of time spent shopping; the last column converts this back to seconds. We thus see that browsers spend on average 2.576% more shopping than do goal directed shoppers on average, a difference which is significantly different.

Next, we will specify uncertainty as High vs. Low instead of country of origin. The R2 is still 0.142 indicating a similar fit to the data. The results follow:
**Dependent variable= LN(Time Spent Shopping)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant***</td>
<td>4.599</td>
<td>0.030</td>
</tr>
<tr>
<td>Browsing***</td>
<td>2.614</td>
<td>0.087</td>
</tr>
<tr>
<td>Low UA</td>
<td>0.009</td>
<td>0.047</td>
</tr>
<tr>
<td>High UA</td>
<td>0.039</td>
<td>0.035</td>
</tr>
<tr>
<td>Low UA * Browse</td>
<td>0.005</td>
<td>0.141</td>
</tr>
<tr>
<td>High UA * Browse</td>
<td>-0.037</td>
<td>0.103</td>
</tr>
</tbody>
</table>

*** indicates significance at 1% CI

I observe that the results are fairly similar. Browsing is still the only significant predictor of time spent shopping, and we see that browsers now spend 2.614% more time per store than do goal directed shoppers (instead of 2.576%). Just like for model II, I exclude all Portuguese phones, and rerun both models. The results hardly change. We can thus say that these results are not the result of a misspecification.

I characterize browsers as shoppers who are in an information seeking mode and generally enjoy the act of shopping. Academics argue an inverse relationship between knowledge and search time (Anderson, Engledow, & Becker, 1979; Moore & Lehmann, 1980; Newman & Staelin, 1971; Swan, 1969); thus consumers who are in a knowledge acquisition stage will spend more time doing so. As a result, browsers, who seek information, will spend more time in a store due to the negative relationship between knowledge and time spent. Similarly, consumers who browse because they enjoy shopping, will have a higher motivation to spend time than those who don’t (Schmidt & Spreng, 1996).

In line with these earlier findings, I also find clear evidence in favor of $H_{2B}$ which postulates that browsers spend significantly more time per store than goal directed shoppers at 1% CI. Similar to previous hypotheses involving culture, we do not find support for $H_{3B}$, which postulates that high uncertainty avoiding countries exhibiting
characteristics of a goal directed shopping mode will spend less time shopping per store relative. The reasons for this are similar to those for model I and II and are discussed in the next section.

**Uncertainty Avoidance: No Predictive Power in our Studies**

Uncertainty avoidance is not a significant predictor in any of our models. There could be any number of reasons for this.

The most logical explanation is that we do not measure a consumer’s uncertainty avoidance directly; we merely note where their cell phone originates. Thus, a consumer with an English cell phone could have a totally different nationality and thus a different uncertainty avoidance index. The biggest contributors to this noise will be (i) expatriates living and working in Portugal and (ii) Portuguese expatriates working in one of the other countries in my data (France, Germany, UK and Spain). These (expats) will have a phone number which doesn’t match their nationality. However, the assumption here is that many of these expats – especially those who have been residing for a long time in a country different from their own – will already have acquired at least some cultural traits from their country of residence. This is a very reasonable assumption as a very often cited consequence of emigration is the experience of acculturation by expats (Sam & Berry, 2006).

The second reason for our finding is the location of the test environment, namely an international airport. This could be cause for a natural selection and bias the data which would seriously alter my results. This is not a new phenomenon. Hannerz (1990) found that cosmopolitans (i.e. global citizens) do exhibit less cultural differences than pure locals.

Third, regarding nationalities, we only have observations from 5 countries in our sample; if we were to collect data from a greater variety of cultures (>25), we can place greater emphasis on the findings. Other academics agree that when using just one dimension of culture in order to predict consumer behavior, a very large number and especially international group of cultures is required (Sivakumar & Nakata, 2001). Similarly, the differences we do find might be attributed to other characteristics. Are these differences
in uncertainty avoidance large enough to warrant the differences found in shopping behavior? If the variety of culture is not large enough, our findings could be spurious. Contrarily, I argue that although the difference in the dimension of uncertainty avoidance is quite wide (e.g. U.K. vs. Portugal), the difference in culture in general is not large enough to warrant a significant change in behavior. This is even more so affected by the open borders we enjoy in the EU (discussed below).

Finally, some academics have argued that culture is not a driver of consumer behavior because of correlations with other factors such as GDP or because of ecological fallacy (McSweeney, 2002). Other academics have also argued that nations are not the appropriate unit of analysis, and that cultures can exist within regions as well as countries (McSweeney, 2002; Sivakumar & Nakata, 2001). Similarly, Yip (1995) argues that national borders are losing importance as indicator of international activities. This is the result of a multitude of events including globalization, open borders and economies, global media, information transmission and much more (Steenkamp & Ter Hofstede, 2002; Yip, 1995). Thus, we might conclude that since the introduction of the Schengen Agreement in 1984, countries within the EU (especially core members such as Germany, Netherlands, and France) have cultivated their own culture, mitigating the effects of their own national culture.
Overview of Findings

**H₁** Consumers from high uncertainty avoiding countries will more likely be browsing than in a goal directed shopping mode compared to people from less uncertainty avoiding country.

**Findings** I find evidence to support that country of origin has an effect on a person’s shopping mode; however the relationship is not as was expected. English consumers are most likely to browse, followed by Germany and Portugal respectively. French origins have no impact.

**H₂A** Consumers in a browsing shopping mode will exhibit the highest purchase incidence compared with consumers in a goal directed shopping mode.

**Findings** I found strong support to indicate that browsers are more likely to purchase than goal directed shoppers. Browsers are 2.29 times more likely to make a purchase compared to goal directed shoppers.

**H₂B** Consumers in a browsing shopping mode will spend (on average) the most amount of time shopping per store compared to consumers in a goal directed shopping mode.

**Findings** I find strong support for H₂B; browsers spend (on average 2.576%) more time per store compared to goal directed shoppers.

**H₃A** Consumers from high uncertainty avoiding countries exhibiting characteristics of a goal directed shopping mode will have a higher purchase incidence relative to consumers from uncertainty accepting countries.

**Findings** I find no evidence to support hypothesis H₃A.

**H₃B** Consumers from high uncertainty avoiding countries exhibiting characteristics of a goal directed shopping mode will (on average) spend less time shopping per store relative to consumers from uncertainty accepting countries.

**Findings** I find no evidence to support hypothesis H₃B.

5. **Conclusions**

The aim of my thesis is to investigate the effects of certain consumer characteristics on their shopping behavior. I investigated the impact of origin on shopping mode, and the impact of origin and shopping mode on purchase incidence and time spent shopping.

I did this by using data collected at the duty free shopping area of a mid-sized European airport serving a metropolitan area of slightly more than 2 million people. The data
covered shopping behavior at four different stores in the duty free area of the airport. I have data regarding whether a consumer entered a specific store, how long she stayed there, whether she made a purchase or not. I also know how many stores a consumer visited, and where the consumers’ cellphone originate. Using this data I made some interesting discoveries.

First, consumer culture has a significant effect on shopping mode, although the relation is not the same as I originally hypothesized. Consumers from the UK are roughly 26% more likely to be in a browsing stage than Germans and the Portuguese; the Spanish consumers even less inclined to browse. German, Portuguese and Spanish consumers are all roughly equally likely to browse. Consumers from the UK are roughly 21% more likely to browse than the French at 10% confidence interval. The unexpected sign of these effects (together with the fact that prior researchers have found that demographics are unable to explain the drivers of shopping mode) suggests that more research on this topic would certainly be useful.

Second, I learned that consumers in a browsing shopping mode (i.e. they visited all 4 stores) are significantly more likely to make a purchase than consumers in a goal directed shopping mode (i.e. visited only 1 store). Specifically, consumers in a browsing mode were 2.29 times more likely to make a purchase than goal directed shoppers. These consumers also spend significantly more time shopping (on average, per store) than their goal directed counterparts. Browsers spend 2.576% more time per store than do their counterparts. This coincides with previous research; however my results are quantifiable because they are based on the real situation instead of self-reported values.

Third, despite a strong body of theory suggesting that culture would moderate the effect of shopping mode on shopping behavior, I found no significant interaction between culture and shopping mode. This was true both for purchase incidence and for time spent shopping per store. This could be the result of our so called cosmopolitan effect or maybe the effect just does not occur in this setting.
In conclusion, I have shown that consumer browsing behavior is a significant indicator of purchase intention and on time spent shopping.

**Practical Implications**

These findings are especially practical for (local and international) retailers, marketers, and marketing intelligence providers.

**Marketers in General**

It is important for (inter)national marketers to recognize that culture and/or country of origin does impact the way consumers perceive products and services. Consumer wants are not endlessly converging and neither is our individual values disappearing. While some cultures prefer to browse more, other cultures prefer to get in and get out. Marketers need to realize this and choose which way they will want to position their brands. After making this choice, it is especially important for marketers to communicate this position, especially in countries where it matters. This is especially important for multi-national corporations operating across many different borders. Similarly, marketers need to pay extra attention as to not make rash stereotypes or oversimplifications of their customers.

**Local and International Retailers**

Knowing that browsers are more likely to make a purchase, and that they spend more time in-store is helpful in training personnel. If personnel can identify browsers based on shopping time, they can anticipate a purchase and spend more time assisting the customer. Shopping malls could develop smartphone apps which track consumer movement, and could relay real-time information about the consumer’s behavior to the individual retailers. They could include such a function in a customer loyalty or special offerings application. An alternative method for this could include attaching RFID (transmitting) chips to shopping carts in order to track a consumer’s movement within the same shopping mall and relaying this information to individual retailers.

Knowing that certain cultures are more likely to browse than others is a major insight for local and international retailers. With this knowledge, retailers can alter their stores in order to satisfy local preferences. An international retailer could adjust their stores in
the UK to accommodate browsers, whereas the same retailer would adjust their Portuguese stores to more goal directed consumers. For example, the Apple store in London would have more products on display, whereas the Apple store in Porto would be design for people who want easy access to products and they want it quick. Similarly, retailers should try to understand the intentions and interests (browsing or goal directed) of their current customers in order to better serve them.

**Marketing Intelligence Providers**

Having seen the power of information, it is important to realize that accurate and concise information regarding consumer shopping behavior is a very valuable commodity. It is therefore important for providers of marketing intelligence to realize that information regarding consumer shopping behavior is of crucial importance in order to help serve the consumer. It is also important for them to realize the importance of accurate and complete information. Therefore, our insights should help providers of marketing intelligence continue collecting information and investigating the effects of culture on shopping behavior. I also recommend finding out a way in which nationality (and other customer demographics) can be more accurately determined in real time, so that retailers can push specific adverts to a customer either through MMS or a specific shopping mall application that could simultaneously function as a customer loyalty card.

**Limitations**

I pro-actively look at two sources of limitations: culture as a predictor of behavior, and our data.

**Limitations of Culture**

Although a lot of evidence has been found to suggest national culture is a strong driving force behind consumer decision making, there is also evidence to suggest this relationship is attributable to something else. Even Geert Hofstede himself is quick to point out flaws in his own work. The seemingly biggest critic of his work is Brendan McSweeney. In his widely cited (850 cites since 2002) article “Hofstede’s Model of National Cultural Differences and their Consequences: A Triumph of Faith – a Failure of Analysis,” McSweeney highlights a number of a major flaws. We will analyze the most prominent arguments and look for counter-arguments.
Possibly the largest limitation of said cultural dimensions is the way that the dimensions have been recorded (McSweeney, 2002). Geert Hofstede collected data from 117,000 IBM employees in order to build his dimensions of culture. The biggest criticism stems from the fact that large firms can usually characterized by their own form of culture. IBM's culture for example was known for its white collar shirts, administrative assistants, and its individualism (Lagace, 2002). This individualism is what drove IBM sales in the '60's and 70's, however this was not tenable anymore starting in the 1980’s and 90’s. IBM's culture went through a transformation, towards a more collectivist culture (Lagace, 2002). It is therefore highly possible that the type of people IBM attracts through its own corporate culture are different from others with the same nationality. Especially considering IBM’s culture as being described as individualistic; this could adversely affect the reliability of Hofstede’s dimensions of culture.

In response to this piece of criticism, Hofstede (2002) elaborates that his survey measures merely the differences between (national) cultures; Hofstede argues that information regarding cultural differences can be provided by any set of practically similar data sample.

Another fair point of criticism is the fact that some dimensions of culture are positively correlated to GNP per capita (Hofstede, 1984). The idea is that as countries develop their economies, and become wealthier, they also have a strong tendency to become more like other developed countries as a result. In other words, if a country acquires the resources that let people “do their own thing,” they will most like do just that.

Hofstede (1988) responds to this issue by stating that his findings persuasively prove that individualism is a product of wealth, and not vice versa and that this is just a product of cultural differences embedded in hundreds of years of experience.

Furthermore, McSweeney (2002) points out that not only are surveys an unsuitable tool for measuring differences in national cultures, but that nations are not even the most effective units for studying culture. McSweeney (2002) argues that there is absolutely no consensus in the wider literature that the units of analysis should be individual countries. Regarding the tool used for collecting data, McSweeney argues that while
117,000 respondents seems like a lot, a large dataset does not in itself guarantee representative findings. Either way, after taking a closer look at the data reveals that for many countries, the sample size was even considered “miniscule;” in 15 countries, the number of respondents was less than 200, which according to McSweeney is far too little (McSweeney, 2002).

Regarding the units of analysis; Hofstede (2002) acknowledges that it is not ideal to use countries as units of analysis; however, he notes that it is the only unit of analysis available and that it is better than nothing. Regarding the matter of completeness, Hofstede does not reply directly to the critique. However, he makes a note on the extensive validations that other academics have added to the body of literature. He notes that since his publications the world has seen 4 large scale replications (covering between 15 and 32 countries each), 400 independent and significant correlations, and over 1500 cited sources (2002) adding to his findings.

We have reviewed a number of valid points of critique, and provided some counter arguments for them. Whether these ideas are right or wrong is not necessarily the point. The fact that these dimensions of culture have been used so widespread, and by academics in many different fields of study is a testament to the applicability of the five dimensions of culture stipulated by Geert Hofstede.

Limitations of the Data
Although the data has its obvious advantages such as being unobtrusive, accurate, unbiased and complete, there are (as with all data) some limitations as well.

First of all, there are a significant amount of consumers who spend less than 5 seconds in a store and still manage to make a purchase. The dataset records a purchase as a consumer who spends more than 30 seconds near the cash register, thus we would expect that consumers who make a purchase straight away have a dwell time of at least 30 seconds. Second, we may have missed a number of consumers who spend less than the required 30 seconds at the cash register when there are no queues and payment is made quickly (e.g. exact change). Third, the data does not tell us very much about what a person is doing or what he is buying. Similarly, a consumer might spend
30 seconds near the cash register, but may be asking a question or doing something else. Fourth, the data has been collected in an airport with only four stores. So first of all, consumers may otherwise have not visited these stores but because of their lack of better alternatives, shop out of boredom. Also, analyzing an environment with a greater variety of stores may be more revealing. Finally, the data is heteroskedastic, and since PASW does not include White’s Error Terms or Weighted Least Square, we use a confidence interval of 1% as a crude method of error correction. Ideally I would like to use either White’s error terms or WLS.

**Future Research**

There is a lot of room for future improvements. First of all, in order to better inspect cultural differences data collection should be done at an international terminal. Here we could inspect the behavior of people originating from more than just 5 countries. This would provide the analysis with significantly more predictive power. Second, in order to complement the data we have now; it would be very insightful if purchases were tracked as well; either specific products, products classes or purchase amounts could help researchers to new understandings. Besides purchase behavior, other characteristics such as sex, age, and income will help make the dataset even more valuable. Thus, companies need to find a way to retain the automated process of this technology, and complement it with cooperation of retailers, software that empowers the consumer, or employ people on-sight to make observations.
6. Appendix

Data Cleaning
Now that we know which variables we need to use, we can take a closer look at the data. It includes 45,980 observations and a number of variables. The ones we will use are described below:

- X Dwell: this indicates the amount of time a consumer spends in store X
- X Bought: whether a consumer makes a purchase at store X
- Country: country of the network operator of the consumer
- #stores: this is the number of stores a consumer visits

I first start by “eyeballing” the data. This leads us to find some interesting things. First of all, there are no consumers who purchased at more than one store. Secondly, consumers visited either 1 or 4 stores; nobody visited either 2 or 3 stores. Finally, I notice that quite a large number of people spend less than 5 seconds in a store before making a purchase. I also removed all cases of people who did not visit any stores as I am interested in investigating effects on shoppers, not consumers who did not do any shopping. We are left with 22358 cases of preliminary data (before rigorous cleaning).

I notice a number of errors with the data. First of all, the datasheet indicates quite a number of consumers who spend less than 5 seconds in a store but still manages to make a purchase. I also notice that the datasheet does not accurately record the number of stores a consumer visits. Finally, there are 950 cases of missing nationalities which I have removed. Although not necessarily a problem, the time stamps are encoded by excel as a 12-hour clock. Therefore, if a consumer spends 2 minutes and 25 seconds in a store, it is recorded as 12:02.25 AM. In order to fix the problems mentioned above, I use excel to create new variables, convert old ones, and create a number of new variables.
Using the MINUTE and SECOND function in excel, I convert “dwell” times into seconds spent per store.

Similarly, I convert all other variables from time into seconds.

Using the COUNT function, I create a new (accurate) variable in excel which counts the number of stores visit based on dwell times larger than 0.

Created a new variable “purchase incidence” to record whether a consumer has made a purchase or not, irrespective of the store the purchase was made in.

Created a variable to indicate average time spent per shop visited.

After rearranging the data, removing certain observations and creating new variables I proceed to take a closer look at the data to spot outliers. I will focus my efforts on taking a more in-depth look at the data to spot outliers and deal with them appropriately.

An outlier is an observation or group of observations that deviate from other observations in the same data set, and has severe impact on the results of a statistical analysis.

Using a two-step cluster analysis, I identify two distinct groups of individuals in the dataset; one group has a mean shopping time per store of 93 seconds, and the other 502 seconds. Accordingly, I split the group up into consumers who visited 1 store and consumers who visited 4 stores and identify outliers in each specific group. First, I use a cumulative distribution function of the Mahalanobis distance to investigate whether there are any unusual combinations among the dwell times of the four stores. I find only 6 multivariate outliers which differ significantly from the rest and remove these.

Subsequently I use a boxplot to identify whether there are any univariate outliers remaining in the variable time spent shopping per store. I find none, and also notice that the data seems quick nicely distributed.

After selecting only the cases who visit 1 store, I investigate the same parameters. I find only 3 univariate outliers using a boxplot and notice that the data is strongly right skewed. This is not a strange finding though, as we have over 4,000 cases of consumers who only shop for a couple of seconds before making a purchase.

Finally, I am interested in a number of characteristics of the final data. Using a histogram of time spent per store, I can see that the data is not normally distributed, although this is not necessarily a problem as we rely on the law of large numbers (many
observations), it is still wise to keep this in mind. The root cause of this might be the 4,000 observations entering a store for 2-3 seconds and making a purchase. This might suggest a large body of goal directed shoppers in our sample. After removing all of the shoppers who spent less than 5 seconds in a store and made a purchase, the histogram looks much more normal. I observe two peaks, suggesting we might have 2 distinct groups in our sample. I observe a small group of shoppers who spend between 30 and 200 seconds in a store, and a larger group of shoppers who spend more than 200 seconds per store. Using a two-step cluster analysis with a log-likelihood distance measure, I identify two distinct groups of individuals in the dataset; one group has a mean shopping time per store of 93 seconds, and the other 502 seconds.

Thus in short, we start with 45,980 cases, and after removing everyone who did not spend at least 1 second in any given store, we are left with 22,358 cases which seem to be clustered around 2 distinct groups. We create a number of variables for our regressions; the most important being time spent per store. After analyzing boxplots and using Mahalanobis $D^2$ to investigate univariate and multivariate outliers respectively, we remove another 9 cases and are left with 22,349 cases for our final analysis.

Two-step cluster analysis

<table>
<thead>
<tr>
<th>Cluster Distribution</th>
<th>% of Combined</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>51.6%</td>
<td>51.6%</td>
</tr>
<tr>
<td>2</td>
<td>48.4%</td>
<td>48.4%</td>
</tr>
<tr>
<td>Combined</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Total</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>Time per Store (sec)</th>
<th>Mean</th>
<th>Std Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>502.87</td>
<td>155.287</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>93.04</td>
<td>85.115</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>304.31</td>
<td>241.050</td>
<td></td>
</tr>
</tbody>
</table>
Box plot (4 stores visited)

Histogram of Time Spent per Store
Box plot (1 store visited)

Histogram of Purchase Incidence

Purchase Incidence

Frequency

Purchase Incidence
Histogram of Time Per Store (after)

Residual Scatterplot Model
Breusch-Pagan and Koenker Test

Number of predictors (P) 9

Breusch-Pagan test for Heteroscedasticity (CHI-SQUARE df=P) 195.981

Significance level of Chi-square df=P (H0:homoscedasticity) .0000

Koenker test for Heteroscedasticity (CHI-SQUARE df=P) 268.171

Significance level of Chi-square df=P (H0:homoscedasticity) .0000

Histogram of Residuals Model

![Histogram of Residuals Model](image)
7. Bibliography


