PARRISH, JASON S., M.A. Geographic Information Systems and Spatial Analysis of Market Segmentation for Community Banks. (2009)
Directed by Dr. Rick L. Bunch. 79 pp.

The purpose of this research was to determine the value of incorporating consumer behavioral datasets, specifically lifestyle market segmentation, into traditional site selection and location models for community banks. A bank provided the sample of 3,803 customers in 161 Census block groups in its trade area.

The trade area was created using GIS. The necessary customer and lifestyle market segmentation data were also prepared using GIS. Four separate ANOVAs showed that consumer behavioral datasets are significant in predicting market penetration. Market penetration is a critical element of assessing a firm's strength; therefore consumer behavioral datasets are valuable in market analysis and site selection for community banks.

GEOGRAPHIC INFORMATION SYSTEMS AND SPATIAL ANALYSIS OF MARKET SEGMENTATION FOR COMMUNITY BANKS

by

Jason S. Parrish

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Greensboro
2009

Approved by

Rick L. Bunch, PhD

Committee Chair

To my wife, for without her support and understanding this would not be possible. This is as much a product of her sacrifice as it is my work. Also, to my parents, for if not for their constant encouragement, I would not have begun this journey.

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair __________________________________
Rick L. Bunch

Committee Members __________________________________
D. Gordon Bennett

____
Elisabeth S. Nelson

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

ACKNOWLEDGEMENTS

Thanks to Dr. Rick Bunch, Dr. Gordon Bennett, and Dr. Elisabeth Nelson for their expertise, guidance, and open doors. Thanks to Dr. Bunch, Tom and Richard for the opportunity. The research was funded through the Center for Geographic Information Science at The University of North Carolina at Greensboro by an innovative community bank which shall remain anonymous.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

The purpose of this thesis is to determine the value of incorporating consumer behavioral datasets into traditional site selection and location models for community banks. Consumer behavioral data such as lifestyle/life-mode classification should prove to be important in the analysis of prospective markets for community banks. Understanding consumer behavior should also provide financial firms robust means for identifying and targeting individuals that have a high likelihood of becoming customers. Further, the use of spatial analysis can also delineate regions of high consumer demand and expected returns. This study examines the influence of household characteristics on consumer behavior. The focus will be explaining a bank's levels of market penetration across its trade area given a set of lifestyle/life-mode characteristics.

This research addresses a major problem in many marketing campaigns conducted by small community banks. There is an underutilization of consumer behavioral datasets particularly within a spatial context for business and marketing. Lifestyle/life-mode datasets are valuable because they consolidate disparate data into readily accessible information about consumers. Studying spatial patterns and implementing location models offers value for academic research as well as businesses. Businesses, for example, looking for the "perfect" customer can use spatial analysis to identify high concentrations of individuals meeting their selected behavioral criteria. Approaches grounded in consumer behavioral research and geographic space can reduce initial capital outlays often associated with targeted marketing campaigns which ultimately increases return on investment. The goal is to leverage the spatial aspects of Geographic Information Systems (GIS) to develop methods that can be ultimately transferred to consumer behavior research that varies in size, scale and scope. The results of this research show that lifestyle/life-mode data are a valuable part of consumer behavioral research, and should contribute to current spatial procedures used in GIS.

The following represents an outline of this research. Chapter II provides the literature review that focuses on traditional location models and consumer behavioral research in geography. This chapter also provides literature on the use of GIS for site selection and location models. Chapter III outlines the data collection methods and procedures for organizing and preparing data for analyses. It is important to note that in this section all data and geospatial datasets obtained for the community bank under examination in this research will be referred to as “Guilford Savings Operations” or “GSO.” Chapter IV describes several statistical procedures used to analyze the computed market penetration and consumer behavioral data for GSO. The final chapter offers conclusions based on knowledge of GSO’s preferred customers and the characteristics of the total population within their trade area as described by the lifestyle/life-mode data. These conclusions were also supported by the results of the statistical analysis.

CHAPTER II

REVIEW OF THE LITERATURE

A study of the application of locating community bank branches is an excellent opportunity to employ geographic principles. The study is incomplete, however, without understanding the origins and theory behind site selection. The study of locating bank branches has evolved through several stages. Location research began in other industries when investigators such as William Applebaum (1966b) and David Huff (1964) explored the concept of a trade area and how the understanding of a trade area makes it possible to plan new locations. Later, researchers and banking professionals looked into the specific needs of banks planning new locations.

Bank location research has its roots in studies on supermarket and retail location. The early location studies were conducted for companies in these industries. Researchers worked to understand consumer demand, which was often addressed in a spatial context. Of course, consumer demand is not the only factor that determines suitable sites for new store locations. Additional research explained the importance of other factors (e.g., traffic flow, parking availability and competition) that can impact the site selection process.

An understanding of the history of marketing geography, therefore, is important for understanding the theory behind deriving trade areas. Store location research has its origins in marketing geography. Central Place Theory (CPT), for example, was a major concept that shaped the discipline and the concepts of trade areas. Later on, key players such Huff and Applebaum drew from CPT to shape and provide meaning for their own versions of trade area research. Each derived theoretical trade areas through different methods, but each version was meant to incorporate the spatial aspects of consumer demand. Both methods allow consumer behavior to be visually mapped. Two important functions of trade areas are 1) to improve existing markets and 2) to predict the success of future markets. This study will focus on function number 2, predicting the success of future markets.

The two most influential methods of determining store trade areas are gravity models and analog models. Huff worked with gravity models, which are empirically derived measurements of consumer demand. Gravity models, are very similar to Isaac Newton's gravitational law and help account for human behaviors related to shopping activities (Lee and Pace 2005). In the Huff Model, once consumer demand is determined, it can be traced using demand contour lines and visualized in map form or represented as demand zones (Figure 2.1). The analog model, on the other hand, requires detailed research of an existing market area. The ideal situation for analog models occurs when the trade area of an existing store location is fully understood and this understanding is used to find analogous characteristics of other geographic locations. The purpose is to identify new and successful markets not already included in the store's trade area (Applebaum 1965).

In comparison to retail shopping centers there are different factors to consider when planning bank branch locations. Retail firms and supermarkets, for example, often take floor space into account to determine the demand created for a good. The most important factor for bank branches, on the other hand, is deposit potential. A bank

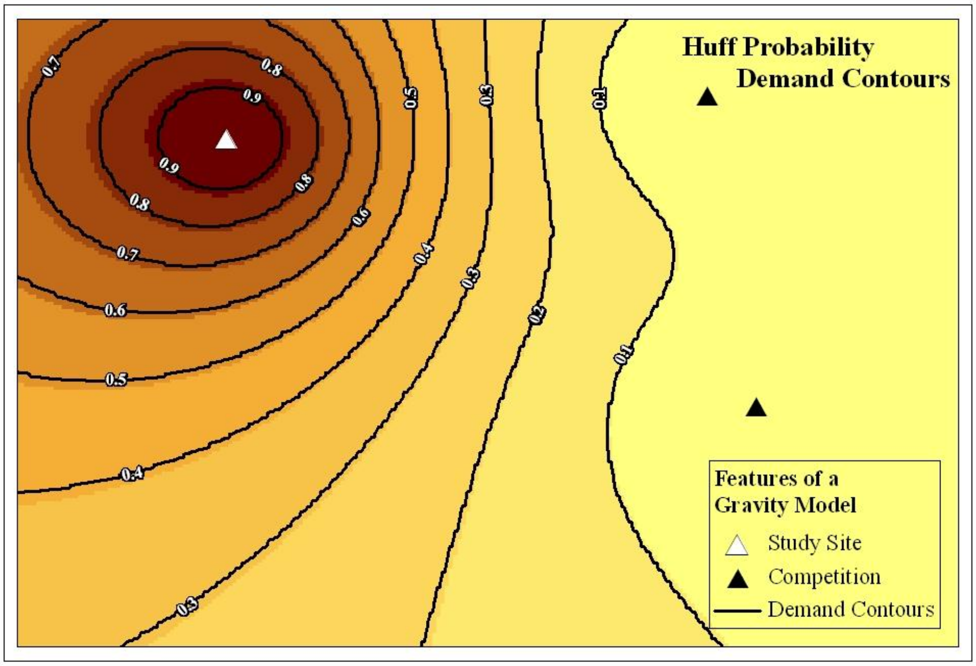

Figure 2.1. Huff Probability Demand Contours

generates its profits on interest received from loans, but it cannot loan money without adequate deposits to create a pool of money from which to lend. Competition from other banks is an important consideration as are the demographics of the market.

Because this study aims to incorporate lifestyle/life-mode groups into the site selection process, the history and theory of this type of marketing research merits attention. Additionally, lifestyle/life-mode data are collected by enumeration units such as block groups that are ideally suited for analysis in GIS. Using GIS and these data together can define current trade areas as well as identify new ones.

2.1 Marketing Geography

Marketing geography provides the most basic approach to store location strategy (Applebaum 1965). Before the 1950s, economic geography was focused on production and transportation of goods to distributors. Applebaum (1954) argued that the function of marketing was the transfer of goods from distributors to consumers. He believed marketing had been overlooked by the field of economic geography. Applebaum proposed that marketing geography be concerned with the identification and measurement of markets and networks through which goods are transferred to consumers.

Throughout the next decade, marketing geography became more focused on measuring market areas of existing stores and estimating market areas of potential store locations (Berry 1967). Until this time, store location research had been more of an art than a science (Cohen and Applebaum 1960). The solution was incorporating more facts into the process to allow for intelligent decisions (Berry 1967).

Jones and Simmons (1987), for example, related location analysis to marketing. Marketing is concerned with what is known as the "seven P's: people, product, price, place, physical distribution, promotion, and packaging." Location analysis is often associated with three of these P's: people, place, and physical distribution. Physical distribution refers to location of customers and the connections between geography of production and patterns of consumption (Berry 1967; Jones and Simmons 1987).

2.1.1 Central Place Theory

In the first half of the twentieth century, marketing science was emerging in the academic environments of business and economics, but it was largely ignored by geography (Berry 1967). A notable exception was the German geographer, Walter Christaller, and his colleague, economist August Lösch. They provided fundamental theoretical concepts of Central Place Theory (Brown 1992). The theory states that the most efficient place to locate a new store is exactly between existing stores. By following this pattern, all trade areas will become equal in size and shape (Berry 1967).

The core of CPT is actually based on several implausible assumptions that were not valuable for empirical studies (Stevens 1985). For instance, CPT assumes a microeconomic world in which all consumers are equally wealthy, rational, and evenly distributed through space. CPT also argues the cost of entry into the market is equal for businesses, and that travel costs are the same for all consumers. A final assumption of CPT is that consumers only purchase one product (Brown 1992; Berry 1967; Kivell and Shaw 1984). In this environment, a trade area for a store is a perfect circle with a radius extending from the store and ending where customers realize it is cheaper to travel elsewhere for the product. Demand reaches zero at the point to where consumers are unwilling to travel to purchase the product because the total cost (product price plus travel cost) is more affordable elsewhere. The trade areas ultimately take on a hexagonal shape to include customers that are contained within two overlapping circles (Brown 1992). Removing the competition between overlapping circles is most advantageous for the firms (Scott 1970).

According to Scott (1970), several limitations of central place theory make it difficult for predicting actual trade areas. Beyond the implausible assumptions listed above, central place theory is static, deterministic, and retrospective. CPT does not take into account social or political factors that influence consumer behavior. CPT also has a distorted view on the concept of production. In the real world, for example, there are multiple goods which may be priced differently for various reasons that CPT cannot take into account.

2.1.2 Trading Areas

One of the most important ideas in marketing geography is a trade area. Prior to the work of Huff and Applebaum, the American Marketing Association defined a trade area as "a district whose size is usually determined by the boundaries within which it is economical in terms of volume and cost for a marketing unit to sell and/or deliver a good or service (Huff 1964: pg 38)." This definition was not very specific and was also flawed as it assumed nothing about demand and implied a firm could determine its own trade area.

Huff (1964) made several conclusions about the nature of trade areas as defined by his gravity model. For one, he assumed there was demand for whatever goods the firm was offering. He represented his trade area as a demand surface, and like a topographic map's contours of elevation, he visualized his demand surface using contours. The focus is on a distribution center that may be a firm or an agglomeration of firms (Figure 2.1). The contours represent the boundaries of demand zones reflecting sales potential.

The zones are probabilistic in nature, and the values range from one to zero (Fig 2.1). The total number of potential customers in the trade area is equal to the sum of the expected number of consumers in each of the demand zones. Zones of competing firms overlap resulting in spatial competitive equilibrium where overlapping areas have equal probability. After synthesizing his conclusions, Huff (1964: pg 38) defined a trade area as “A geographically delineated region, containing potential customers for whom there exists a probability greater than zero of their purchasing a given class of products or services offered for sale by a particular firm or by a particular agglomeration of firms.”

According to Berry (1967), on the other hand, an ideal theoretical trade area is a perfect circle drawn around the store (Figure 2.2). It represents the distance at which consumers are no longer willing to travel to purchase a good. Demand for a good is calculated as the price consumers are willing to pay for the good plus the cost of traveling. At a certain point, consumers are no longer willing to travel resulting in zero demand. The distance from a store to where demand is equal to zero is used to compute the radius of the circle. Once the area is found, the total quantity of the demanded product or service can be derived by multiplying the area of the circle by population density.

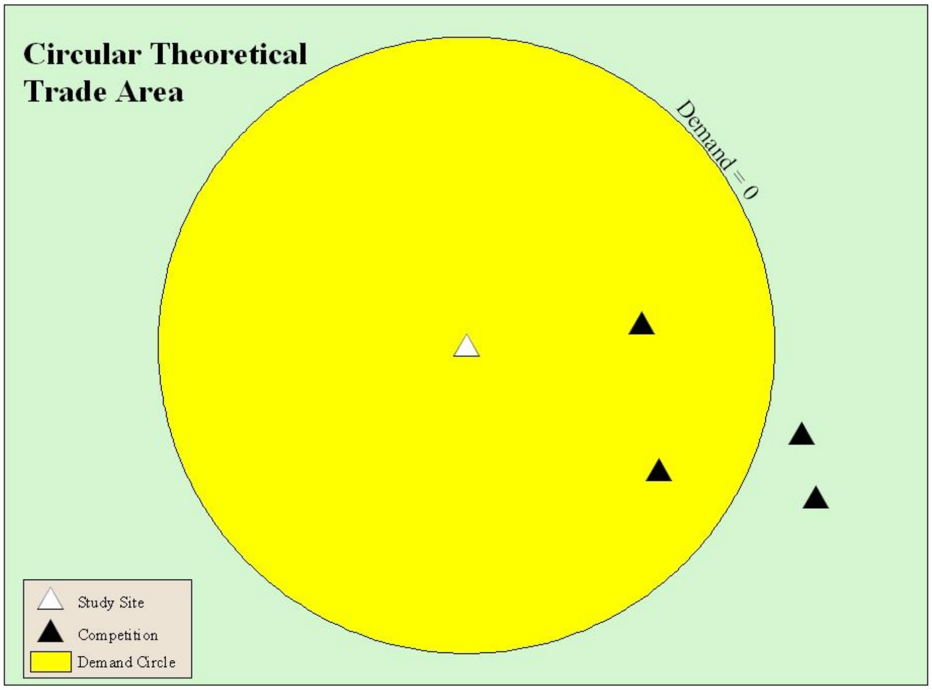

Figure 2.2. Circular Theoretical Trade Area

Berry (1967) used this theory to create a demand cone. In the center of the demand circle, above the firm, a peak represented the demand for the product sold by the firm at a given price. Thus, as travel time increases away from the origin, demand decreases linearly. Kivell and Shaw (1984) noted that in central place theory, demand is equal in all locations within a store's trade area, resulting in a demand cylinder. In gravity models, demand decreases similar to the basic demand cone, but not in a linear manner.

As a result of the earliest trade area studies, several important irregularities are known to exist. First, the proportion of customers patronizing a given location varies with distance from the location. Second, the proportion of customers also varies according to the variety of merchandise (services) carried at the location. Also, the distance that consumers are willing to travel varies according to the types of services. Lastly, a firm's trade area is influenced by the proximity of competition (Huff 1964).

According to Edward Blair (1983), trade area maps are helpful marketing tools for retail merchants because they are useful visual aids for promotional and location decisions. They not only show currently served areas, but also display potential new locations. Trade area maps can also show regions that are suitable locations for targeted mailings or advertising.

Before the use of GIS, a store's trade area could be determined from what is called "customer spotting" (Figure 2.3). This concept involves surveying customers to obtain home addresses or points of origin that are then represented on a map. In the most basic method of representing a trade area, concentric circles can be used to tally customers and sales. Figure 2.4 illustrates Applebaum's (1966a) concentric circles. As an example, if 20 percent of a store's spotted customers come from any zone and the store has $30,000 sales each week, then $6,000 sales are assumed to come from this zone.

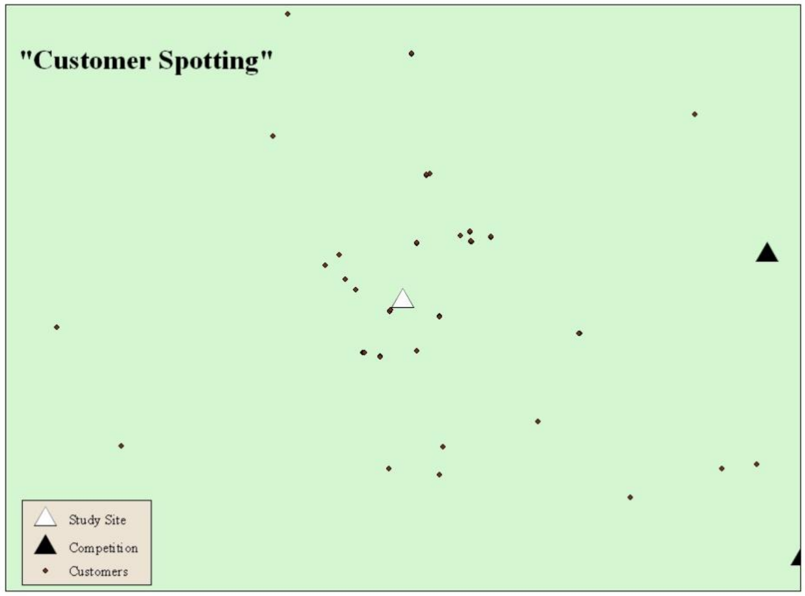

Figure 2.3. “Customer Spotting”

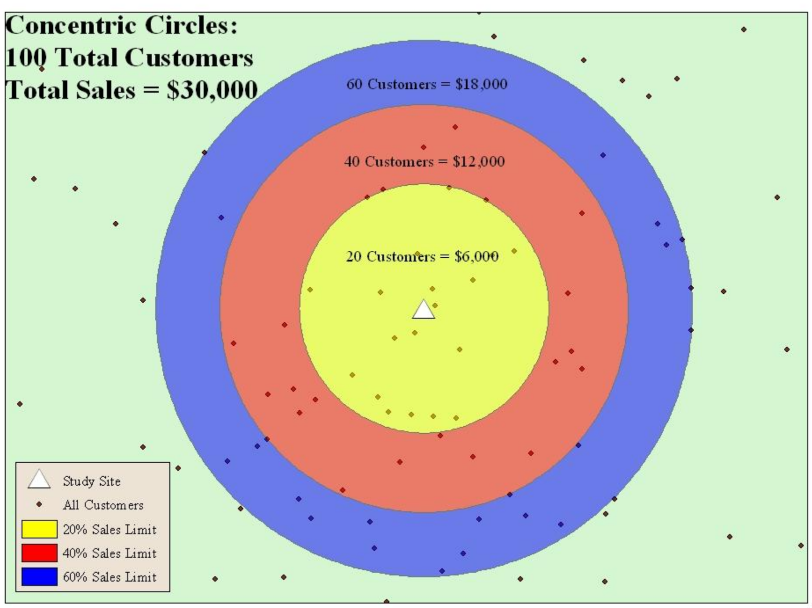

Figure 2.4. Concentric Circles: 100 Total Customers Total Sales = $30,000

Customer spotting is not an exact method and unknown biases can be introduced (Blair 1983). There are three issues that create problems in customer spotting. First, a customer visiting the store during the sampling process may not be a regular or loyal customer. Second, a store with multiple entrances may make it difficult to ensure the entire trade area is equally represented in the sampling. If customers are interviewed exiting the western side of the store only, then many customers who enter on the eastern side may be overlooked, which will erroneously skew the trade area to the west. A final issue surrounds the notion that the timing of the sampling may interfere with the accuracy of how the trade area is depicted. Sampling may only be conducted for a certain period that does not take into account traffic patterns and daily or seasonal customer volume fluctuations. Fortunately, this study did not experience any problems with sampling procedures because all of GSO's existing customers were included from an existing database.

Various approaches have analyzed and predicted retail patronage (Black et. al. 1985). The study by Black, et al. (1985), for example, focused on automobile dealerships, while others have focused on retail shops or supermarkets. There is literature on analyzing retail trade areas, but literature on banking is limited. The development of GIS has allowed the automation of calculating and analyzing trade areas. One major concept is the introduction of drive time analysis (Figure 2.5). John Freehling (1993) proposed using drive time polygons to define trade areas as opposed to concentric circles or locating analogs. His method incorporates distance as a means to segment trade areas to produce drive time polygons based on transportation networks. Drive time analyses do not take into consideration consumer characteristics or the underlying number of potential customers. Drive time analysis also fails to consider competition, resulting in polygons that do not eliminate uninhabited territory such as lakes and airports.

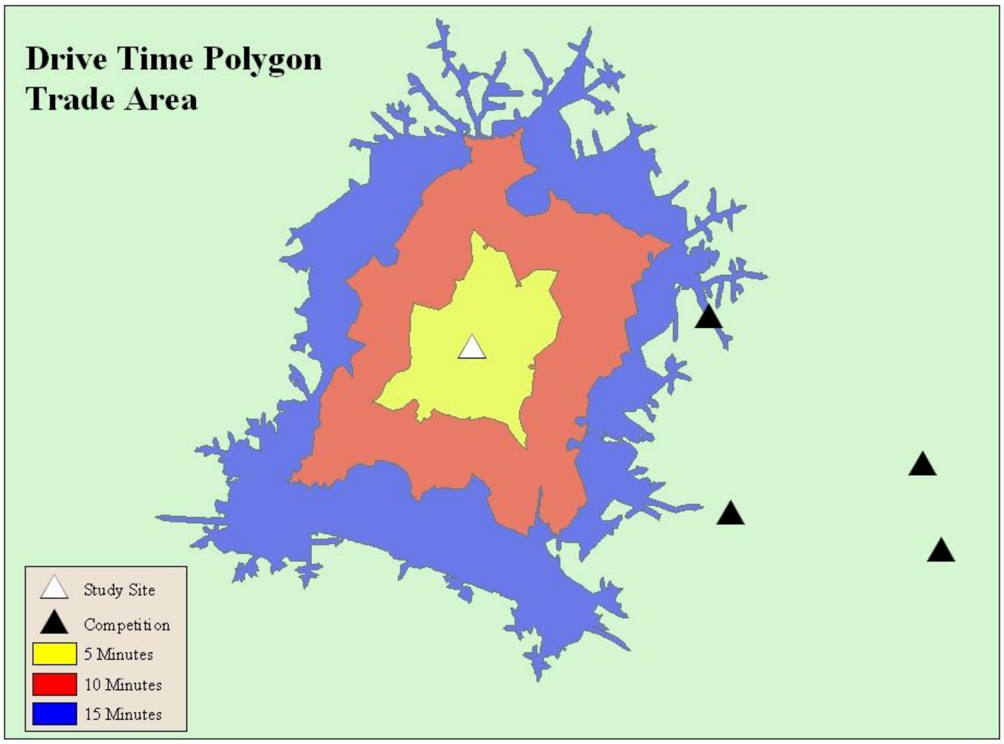

Figure 2.5. Drive Time Polygon Trade Area

2.2 Site Selection Research

Almost a century ago, retail chains were the first to attempt market research with the objective of identifying sites for new stores. The objective was to compare sites within the same geographic business area to determine relative values. The earliest studies focused heavily on pedestrians; however, gas stations began to use vehicle traffic in location studies as automobiles became ubiquitous. Statistics and economics played an important role in empirical studies intended to define trade areas and a store's market share. Later on, developers, investors, and insurers desired an understanding of sales potential around a location and the knowledge of how competition might affect the new site (Applebaum 1965).

Store location research involves the estimation of potential sales within a reasonable error and the computation of a probability of success at each site. A more difficult goal in store location research is the selection of the best location from multiple choices that will provide a maximum market share while minimizing the risk of sales erosion (Applebaum 1965). Predicting sales is possible by comparing potential sites to the performance of existing stores, but only if the market of the existing store is fully understood. If the existing store has a market that is analogous to the proposed site, then sales are likely to be analogous.

Models for selecting store locations should begin with a regional economic analysis. The model designer must also know the business objectives of the firm. As market size increases, the task becomes more difficult. Cohen and Applebaum (1960) detailed several important factors in evaluating sites for large stores. The site must not only be accessible, but it should also be close to population centers or thoroughfares to ensure adequate traffic. Total population, demographic characteristics, growth trends, income, expenditures, and consumer habits are all important details in location studies. Economic stability as measured by economic trends, land values, and employment statistics are also fundamental in the site selection process. These economic measurements help identify the spending habits of the population living in a region, although all consumers may not be equal in their purchasing power. An assessment of competition is also necessary in evaluating store sites. The most useful means of assessing the competition involves estimating the trade areas and market penetration of each competing location. The investigator and firm must finally consider the costs of construction, land, and/or rent.

Applebaum (1966b) further expanded on his earlier work and proposed a detailed process for three types of store location strategies: 1) evaluating a site for building in an area where a firm is already represented, 2) investigating sites in a new market area for a firm, and 3) guidelines to help a firm make an acquisition of an operating location. This study is concerned with the second strategy, investigating a new market area for GSO.

Applebaum (1966b) suggested multiple stages for evaluating a new market area, which do not necessarily have to be performed in any certain order. The firm should first provide the investigator with a specific objective than can be analyzed. This analysis includes an economic base study of the area which involves identifying economic activities, employment, and trends. As part of the methodology, it is required that the investigator obtains a thorough understanding of the underlying population and its characteristics as well as environmental factors such as terrain, infrastructure, zoning, and land use. An inventory and a detailed appraisal of the competition as well as customer attitudes in the proposed market are essential. Finally, Applebaum suggested that the investigator review proposed sales projections to estimate initial capital outlay and return on investment.

More recently, John S. Thompson (1982) provided a basic model for retail site selection and growth that focused specifically on locating new sites. His “pre-strategy” phase identified the locations of ideal customers and emphasized the placement of new sites in those markets. The second phase, called “strategy,” involves dividing the larger markets into submarkets and using these submarkets to conduct detailed site selection using an analog or gravity model.

Peterson and Monroe (2004) expanded Thompson’s (1982) work by providing a more detailed methodology, which divided both of Thompson’s (1982) phases into separate processes. Similar to Thompson’s pre-strategy phase, their first process suggested screening individuals to determine what type of customer is most profitable. In the second process, markets were screened to find locations that are situated among or near the most profitable customers. The strategy phase of Thompson’s work was also divided into separate processes by Peterson and Monroe (2004). Their third process involved an analysis of the regional markets to select the most profitable submarkets. Once submarkets were identified, a fourth process identified specific sites. Peterson and Monroe also required a final process that focused on making a site successful by planning the aesthetic characteristics of the site in relation to its surrounding.

Janet Engle (2007) has provided yet another framework for site selection. She characterized three selection strategies for 1) buying an existing store, 2) buying a franchise, and 3) “starting from scratch.” The present study is interested in the building of a new bank branch "from scratch." Her recommendations for starting from scratch required high traffic volume, growing populations, a strong presence of a targeted population, and a thorough understanding of the competition. Engle recommended collecting data from the Census or independent research firms about the demographics for the proposed sites. The site should also be near other successful businesses some of which must cater to the targeted market.

2.3 Gravity Models

One of the earliest methods for refining trade area definitions was made by David Huff (1963), who believed that some models of determining trade areas were too subjective. He used a mathematical approach that drew from existing gravity models. The retail gravity model was intended to determine trade areas for two competing cities. The gravity models from which he worked included both Reilly (1929) and Converse (1943). William Reilly (1929) created a formula to predict habits of consumers between cities. He designed the model to express the proportion of business each city would attract from the population center of a smaller town located between the two competing cities. Reilly's formula is shown in Equation 2.1.

$$\frac{\mathrm{Ba}}{\mathrm{Bb}} = \left(\frac{\mathrm{Pa}}{\mathrm{Pb}}\right) * \left(\frac{\mathrm{Db}}{\mathrm{Da}}\right)^2$$

Equation 2.1

Where Ba = the proportion of the retail business from an intermediate town attracted by city A;
Bb = the proportion of the retail business from an intermediate town attracted by city B;
Pa = the population of city A;

Pb = the population of city B;
Da = the distance from the intermediate town to city A; and,
Db = the distance from the intermediate town to city B.

The formula reflects distance and population as influential factors. Any small town between two major cities of the same size will be more influenced by the city to which it is closer. Further, proximity is juxtaposed by population size, thus any city with a larger population will have a stronger influence on the small town (Reilly 1929). For a further review of Reilly's model see Scott (1970).

Reilly's gravity model was later modified by P.D. Converse (1943) to identify a breakpoint between the cities. When traveling from one city toward the other, the breakpoint represents the distance between the cities at which the influence of one city ends and the influence of the competing city begins. The breakpoint formula derived by Converse is shown in Equation 2.2.

$$\mathrm{Db} = (\mathrm{Dab})/(1 + \sqrt{\mathrm{Pa}/\mathrm{Pb}})$$

Equation 2.2

Where Db = the breaking point between city A and city B in miles from B;
Dab = the distance separating city A from city B;
Pb = the population of city B; and
Pa = the population of city A.

Beyond identifying breakpoints between cities, these models can define breakpoints between competing firms (Huff 1964). A firm could be surrounded by any number of competitors and its trade area could be defined as the distance between the firm and each of its breakpoints.

Huff (1964) proposed an alternative model that focuses on the consumer rather than the firm. This model described the process by which consumers choose a firm for specific services. The expression is:

$$\mathrm{Pij} = \left(\frac{\mathrm{Sj}}{\mathrm{Tij}^{\lambda}}\right) / \left(\sum_{j=1} \mathrm{Sj}/\mathrm{Tij}^{\lambda}\right)$$

Equation 2.3

Where Pij = the probability of a consumer at a given point of origin i traveling to a particular shopping center j;

Sj = the size of a shopping center j (measured in terms of the square footage of selling area devoted to the sale of a particular class of goods);

Tij = the travel time involved in getting from a consumer's travel base i to a given shopping center j; and

λ = a parameter which is to be estimated empirically to reflect the effect of travel time on various kinds of shopping trips.

One important difference between Huff's approach and Reilly's model is that consumer spatial and non-spatial behavior can be explained by the Huff model (Haines et.al. 1972). The model captures spatial behavior which can be used to predict consumer behavior within the context of multiple store locations. The parameter λ is not considered to be constant and varies according to the type of product. High order goods, such as automobiles, will have a smaller λ than low order goods, such as groceries. Lastly, the model allows the data to be visually represented on a map (Huff 1964).

Later research argued that Huff's model was not accurate in predicting sales for a single location. Studies such as those by Kolter (1971) and Stanley and Sewall (1976), assume that consumers and stores are spatially independent in the Huff model. Lee and Pace (2005) argued modeling spatially independent variables in the Huff model can result in poor estimates and recommended that spatially dependent variables be used.

2.4 Analog Modeling

Another prominent method of estimating trade areas is the analog method. John S. Thompson (1982: pg 113) defines analogs as “snapshot(s) of a store’s performance at a given point in time.” These snapshots are useful for identifying historical trends in the market area for a particular site and predicting future sales. William Applebaum developed the analog model as a means to compare potential sites to existing ones. If there are similarities between the two sites, then forecasts can also be made for the proposed site.

Analog models allow the analyst to predict what may happen at a proposed site, based on knowledge of existing sites (Munroe 2004). To analyze the trade area, it can be subdivided into primary, secondary and tertiary trade areas after performing customer spotting (Applebaum 1966a). If a firm has detailed and extensive records of customers, as is the case with this research, then traditional customer spotting techniques are not necessary (Munroe 2004). The primary trade area is where the store gets the majority of its business. The breakpoint is arbitrary, but Applebaum suggested that 60 to 70 percent of a site’s sales come from the primary trade area. The secondary trade area should account for 15 to 25 percent of sales, and the tertiary trade area represents the remaining customers and sales.

Applebaum provides a step-by-step procedure to divide the trade area. First, a grid of square cells no larger than ¼ x ¼ mi square must be placed over a map of the area, and the population of each must be determined. This grid should be placed over the customer spotting map. The customer density in each square determines estimated sales, if it is assumed that all customers are of equal worth. The estimated sales are subsequently divided by the population to create a sales ratio. A "nearest neighbor" method is applied to select cells (Thompson 1982). The trade area expands outward from the store by first selecting the square closest to the store with the highest sales ratio. Next, the cell with the highest ratios that are either adjacent to the cell that was just selected or to the store is selected. The process is continued until the required number of customers is reached. The number of customers is determined from the 60 to 70 percent range of sales with the assumption that all customers are worth the same amount. A final consideration is to make sure there are no natural topographic or man-made barriers such as bodies of water, mountains, or airports that can affect the practical limits of a trade area (Thompson 1982). The secondary trade area is derived the same way, while the tertiary is extended from the secondary breakpoint to the farthest customer (Applebaum 1966a).

Market share, or market penetration, is measured by dividing the store's sales per capita by per capita sales potential (Applebaum 1966b). Another way to describe market share is the proportion of customers a store has compared to the total population in a given area (Jones and Simmons 1989). Misunderstanding patterns, such as market share, might lead to poor decision making regarding location decisions. A suboptimal location, if selected, may lead to reduced profits, inefficient service, and, in many cases, business failure (Fotheringham 1988).

There are two general ways of determining market share. The first is based on acquiring local knowledge of employees and management, and the second is based on the use of mathematical techniques. The mathematical approach is superior to relying on local knowledge when determining market share for several reasons. First, local knowledge is often uncertain and inaccurate because personal biases may be related to length of employment and the level of experience for different individuals. Second, even if local knowledge were reliable, it does not take into account the understanding of the established relationships between the store and consumption patterns (Fotheringham 1988). Compounding these reasons is the notion that local knowledge is less reliable in the fast pace of modern business (Breheny 1988). New technologies and increasing demand in a consumer-driven society make using mathematical methods quicker and more efficient.

Incorporating census data such as total population, average family size, and median income into Applebaum's (1966b) well-established method can be a very useful approach for determining market share throughout a trade area. In this approach, it is necessary to identify the strength of market penetration so that the census data can be classified by enumeration units for subsequent and meaningful analysis. This analysis would include creating isolines from a grid representing percent of market penetration and ensuring that isolines corresponded to population density and customer distribution. The output from the analysis could be used to estimate potential sales by comparing proposed sites to existing markets. Oftentimes, subjectivity is used to determine the degree of which markets are analogous. The quality and quantity of the data are directly proportional to the accuracy of the comparison.

Regression models can quickly and efficiently replicate the methods of Applebaum's analog models (Jones and Simmons 1989). These models determine which variables are most important in selecting new sites. Relevant variables for consideration in a regression model can include percentage of population in apartments, number of competitors, number of parking spaces, total population, and demographic factors (Jones and Mock 1984). Boufounou (1995) used stepwise multiple regression to determine what variables influenced the success of bank branches in terms of deposits. The important variables derived from his analysis were used to predict the potential deposits of new bank branches. Boufounou found that deposit potential is the primary factor of importance for bank executives in measuring success of current branches and in site selection. This notion was also echoed by the executives at GSO who expressed the desire to increase deposits.

It is important to remember that there is no such thing as a perfect analog (Munroe 2004). An analyst must use his or her judgment or experience to determine what is different between the two markets and assess the importance of these differences. Two analysts with different levels of expertise might delineate different shaped trade areas when studying the same site (Thompson 1982). Many trade areas, for example, will take on an amoeba-like shape, but these shapes may vary greatly depending on individual knowledge of comparing markets. Sales figures are often not explicitly transferred from one branch to another (Munroe 2004). They must be adjusted to account for subtle variations between the markets. The inherent subjectivity of this reasoning is the greatest weakness of analog modeling. Despite this weakness, there are a couple of important strengths associated with the analog method. First, analog modeling is intuitive and simple (Munroe 2004). In addition, subjectivity can actually be a great strength of analog modeling, especially if an analyst leverages his or her experiences to effectively evaluate and forecast potential for new sites.

2.5 Location Analysis for Bank Branches

In the early 1960s, many banks believed opening a new branch in a shopping center would provide the best location (Stahel 1969). Shopping centers, for example, provided an established commercial area, required no need for a property search and made market entry relatively easy. Further, renting shopping center space provided banks with a predictable fixed cost, and parking was widely and readily available. Another benefit of locating in established shopping centers was the reduced cost of advertising due to the fact that customers frequented the shopping center.

Stahel (1969), however, believed that a shopping center was not a good place for a bank. Locating in a shopping center meant the bank would lose its identity and its autonomy. It also lost the business of people who did not want to visit a shopping center while banking. Stahel further argued that a bank would not maximize its potential if it was closed during the extended business hours of the shopping center. The stores around the center may also change over time and attract customers that are not suitable for the bank. Finally, the lack of drive-in facilities could make the bank more inconvenient to customers.

In 1974, Luc Soenen was one of the first to list quantifiable criteria for the decision making process in bank location. His analysis included important demographic criteria such as total family and median household income. Optional variables included per capita income and housing values. Some demographic characteristics can negatively impact bank success. These include small average household size, low percentage married and low percentages of males in the labor force. Other characteristics are positively correlated to bank success, such as high percentages over age 65, increased median education and low percentages of housing units rented.

In addition to the demographic characteristics, Soenen argued that statistics about banking are also required. These requirements include the number of banks in the trade area and the bank-to-population ratio. Furthermore, knowing the number of employees, tellers and customers, as well as the deposit totals of each competitor, would allow the new bank to better understand the market and how to succeed in it. Site attributes, including traffic patterns and general attitudes towards banking would also be very helpful in revealing useful information about the market.

It is important to estimate the deposit potential of a study area after defining the extent of the trade area. Soenen argued regression models are a good approach to estimate potential household deposits using socioeconomic variables. Another more subjective method analyzes the market by the type of deposit by investigating financial statements of local institutions. This method assumes demographics variables to be indicators and drivers of commercial success. Depending on the level of the success of commercial institutions, deposit potential can be inferred. This method is not as precise in estimating deposits among households, but it more accurately identifies commercial and industrial deposit potential.

The final stages of Soenen's process include creating a profit and loss statement, estimating return on investment and making a final decision about a branch location. Soenen believes that a profit and loss statement can be projected for the first ten years of the new branch. It is useful to make sure that returns on investment will be profitable within 3-4 years for the new branch. At this stage in the analysis, it is time to decide whether the location is suitable for a new branch.

Hokey Min (1989) pointed out that previous studies in branch location employed stochastic models to estimate the profitability of potential branch locations. He proposed, instead, a "locational decision support system" that requires two important subcomponents called the model subsystem and the data management subsystem. The model subsystem should be user friendly and easy to understand for the decision maker. The data management subsystem includes external spatial data as well as internal marketing and accounting data.

As Monroe (2004) demonstrated, GIS is well suited for site selection and market analysis. The "locational decision support system" offered by Min (1989), for example, parallels the relational database architecture in GIS as well as other analytical functions. Min's system was designed to allow simultaneous analysis of multiple and quantifiable criteria. These data were required for multivariate analysis to determine the relationship between profitability and data, and to identify population and financial patterns through space with the help of computer-aided mapping. Performing these kinds of analyses and creating relevant maps are common functions of GIS.

The successful execution of a location model for selecting bank branches relies on the quality of location related information (Min 1989). The older models often neglected data management and did not allow for modifications in a dynamic market environment. Min argued that his "Multiple Criteria Decision Making Locational Decision Support System" was flexible, user friendly and accessible and had the ability to process large amounts of data.

2.6 Lifestyle/Life-Mode Segmentation

This section synthesizes some of the previous research on the history and application of lifestyle/life-mode data. Over two decades ago, Kivell and Shaw (1984) clarified that space was not the only variable that could be used to determine potential sales or future success of a firm. They predicted that research would begin to move away from traditional location based approaches toward customer type assessment, formation of store image, and other aspatial variables. Kivell and Shaw's assumption of customer type assessment is strongly reflected in the lifestyle/life-mode segmentation data used in this study. The collection of lifestyle/life-mode segmentation data is based solely on enumeration units that can be analyzed in GIS.

The Community Tapestry (Chapter 4, Appendix A) breaks consumers into lifestyle/life-mode groups. This is often called lifestyle segmentation (Fetto 2004) or market segmentation (Smith 1956). The concept behind segmentation is that a regional market is heterogeneous, but within a market, there are smaller homogenous submarkets (Vyncke 2002).

Jonathan Robbin, who was trained in sociology and statistics, is usually credited with inventing market segments (Goss 1995). Assigning names to the segments was first introduced in 1970 (Fetto 2004), allowing the segments to seemingly come to life. Naming the segments helped businesses to make decisions more easily based on the information. Thompson (1982) said, however, that using market segments had been relegated to a secondary position when analyzing site selection because, at the time he was writing, it was difficult "to ascertain that information on a trade area basis."

The research into the process of lifestyle segmentation, called "psychographics" (Vyncke 2002) or "geodemographics" (Goss 1995), incorporates surveys on activities, interests, and opinions, as well as demographic characteristics. The results are translated into useful lifestyle topologies through the use of cluster analysis. Lifestyles are important to understand from a marketing perspective because they influence not only consumer behavior patterns, but also the processing of various types of marketing (Vyncke 2002). Understanding these patterns may help a business identify a new location by creating an easy way to identify targeted customers as well as add other datasets to an analog model. If an existing trade area can be defined by its lifestyle segmentation, then a region with an analogous lifestyle pattern might be a logical place to begin the site selection process.

Customer profiling is the process of identifying and counting current customers for a firm based on characteristics such as demographics or those found within a psychographic system (Tedrow 2004). It may be meaningful to analyze the ratio of customers to the base. The entire population of the market area represents the base. The proportion of customers within each lifestyle/life-mode group will likely be different. The investigator can better understand which segments might need more attention in a new market. It can also allow the investigator to know which segments are not profitable and should be avoided.

GIS is ideal for analyzing geodemographics for marketing. It makes it easier to use this data at the trade area level, eliminating Thompson's (1982) concerns. Because these consumption and processing patterns are spatially dependent, they can be visually represented in a geospatial dataset or on a map using GIS (Goss 1995). Also, addresses can be spatially referenced using the geocoding process in a GIS. These point locations can then be spatially joined with the lifestyle segmentation geospatial dataset. Marketers have used this method to locate potential clients for stores and for targeted mailing campaigns, as well as for customer tracking. The earliest approaches used Census Tracts as spatial boundaries to divide the lifestyle groups (Fetto 2004). The Community Tapestry dataset is available at the block group level, which provides higher spatial resolution and a better understanding of the variations within markets of varying scales.

It is important to note that lifestyle/life-mode geodemographic analyses are not always completely accurate. Problems arise with geocoding because addresses are not formatted the same way due to data entry errors and varying address styles (Goss 1995). Census data collected at specific enumeration units, such as census block groups, are another problem. Block groups are intended to be homogeneous in terms of population; however, in reality lifestyle variations may exist within a block group. Lifestyle/life-mode datasets are aggregate representations of the average of households within a certain geographic boundary. When a census tract is broken into block groups, the smaller block groups could each be a different lifestyle/life-mode segment. The principle is the same all the way down to the individual address level. It is impractical to believe that every address in the block group is from the same tapestry segment.

It would be extremely beneficial if the Block Group data could be further reduced to the Census Block level. This was possible in a limited way prior to the 2000 Census, because before then, housing value data were also available at the block level. This variable could be used as a surrogate for income or socioeconomic level if it were it still available.

2.7 Summary

This study seeks to incorporate lifestyle/life-mode data into the site selection process. Site selection studies emerged out of central place theory. The crux of site selection is the estimation of trade areas. There are several simple ways to do this, but more complex methods appear to be more accurate. Gravity models select a site and then analyze the market around it. Analog modeling analyzes markets and then selects a site in a similar market. Regression may be used to predict deposit potential and reinforce trade area analysis. Incorporating lifestyle/life-mode data is intended to be introduced early in studies if researches expect to analyze markets. Once an analogous market region is found, specific locations can be selected and deposit potential and other industry specific variables can be estimated.

CHAPTER III

METHODOLOGY

3.1 Study Area

The customer base of the bank in question, GSO, defines the study area. The study area is so defined because the trade is represented as a proportion of their customers. The extent of the trade area limits the area of the analysis. The trade area created for the three longest established GSO's branches extends as far north as southwestern Rockingham County, which is adjacent to Guilford County. To the west, it spills into eastern Forsyth County as far as the town of Kernersville, also adjacent to Guilford County. To the south and east, it is contained within Guilford County (Figure 3.1). Census block groups contained within the trade area in these three counties form the basis for subsequent analysis.

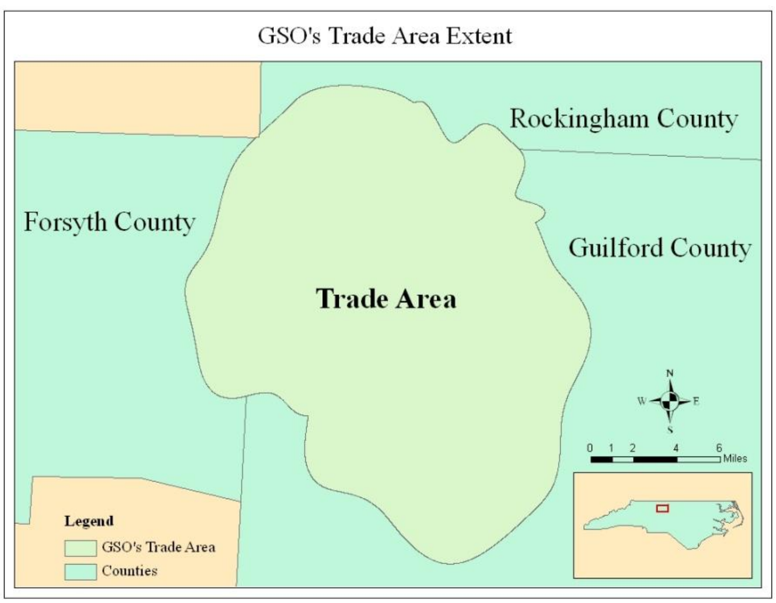

Figure 3.1. GSO's Trade Area Extent

GSO has several branches in Guilford County, North Carolina. Specific locations are withheld, but only three of the branches were used in the study. After constructing trade areas for each branch, it was apparent that some of the newest branches had trade areas that had yet to take on realistic shapes. These branches and their customers, thus, were excluded from further consideration. The total number of remaining customers was 7,296. The remaining branches will be referred to as Branches A-C in order of their establishment. The trade areas of these branches overlap and provide a picture of the GSO's entire trade area.

Recently, the management team and board members at GSO deliberated on the following three questions:

> Do you want your next branch to be located in Guilford County? If not, what counties would you consider?
> Do you want your next branch to focus primarily on commercial or personal accounts?
> Do you want your next branch to have a trade area that overlaps your current trade areas, or do you want it to be separate with a gap between the trade areas?

Results of the inquiry suggest GSO is interested in locations for potential new branches in eastern Forsyth County, Alamance County and southwest Guilford County, all located roughly adjacent to their current trade area. Specific cities of interest are Winston-Salem, Kernersville, Burlington, and High Point. It might be most beneficial to get outside the current trade area for the bank, which is mainly central and northwest Guilford County, including Greensboro and smaller incorporated towns in rural parts of the county.

3.2 Data Preparation

This study followed Applebaum's (1966b) model rather than Huff's (1964), because Huff's model is targeted more toward retail and requires floor space as one of its parameters. Following Applebaum then, existing markets were first spatially defined to draw comparisons. GSO provided a complete listing of their clients' addresses. All of their branches are inside Guilford County. To locate their market, geocoding was used to spatially reference branch and customer locations into geospatial datasets. Geocoding required a table listing customer addresses, which were matched with corresponding street address locations using a GIS. Each branch's customers were geocoded separately to create trade areas for each branch. This represents the customer spotting stage of Applebaum's model. All of GSO's customer accounts were used in the data preparation. Through a geocoding process, 80-85% of each branch's customers were accurately identified and located. Overall, there were 8,753 matched addresses among the five branches. The next step was eliminating duplicate account holders. If a person had multiple accounts, additional accounts created a false impression that many more customers do business with the bank when market penetration was calculated. Once the duplicates were deleted, 5,003 customers remained for use in the analysis.

Creating a trade area in Business Analyst is an automated procedure in which the user selects the parameters that define the trade area. The first step included selecting the layer that contained the branch and the layer that contained the customers. Following this, the exact number of customers defining the trade area was determined. This process involved selecting a number of customers based on a ceiling as defined by a sum of the weighted percentage of each customer (Table 3.1). The literature suggests that all customers are weighted equally when defining a trade area. Branch A had 2,554 customers available to derive its trade area. To establish equal weight, a field was added to the attribute table and each record was assigned the calculated value of 1/2,554 = 0.000391543. Table 3.1 illustrates the results for each branch as well as the number of customers defining each branch's primary trade area.

Table 3.1. Customer Totals for Deriving Trade Areas

Branch	Customers	70%	Weighted Value
A	2,554	1,788	0.000391543
B	1,117	782	0.000895255
C	1,272	891	0.000786164
ABC	4,943	3,461	

Business Analyst was used to select the percentage of customers to define the trade area. According to Applebaum's guidelines, the primary trade area should include 60-70% of the customers for a site. The 70% level was selected to include as many customers as possible and to increase the size of the trade area. This created a larger sample size of customers and total households which lead to a more accurate analysis than a lower percentage of customers.

The last two steps defined the physical properties of the trade area. The center of the trade area could be either the location of the store or the mean center of the store's customers. The latter is the more appropriate option because it is more flexible, accurate, and inclusive of customers. Finally, the desired shape of the trade area was specified. Options are simple generalized, amoeba-shaped areas, or a detailed method that accounts for customer distributions. The first two options only grow from the center until the specified number of customers is selected. The final option was most appropriate for this study because it takes into account customer distribution and is the most accurate. Once this process was completed for Branch A, it was repeated for Branches B and C.

In addition to the locations of the branches and the customers, a second dataset, the lifestyle/life-mode market segmentation set known as Community Tapestry (Appendix A) was also required for the planned analyses. The Community Tapestry dataset, as described in Chapter IV, section 1, is available at every Census enumeration unit with the exception of blocks. Smaller units allow for a more accurate analysis because they represent the most homogenous concentrations of customers. A census tract, for example, might be defined as “Aspiring Young Families” because 60% of its population can be categorized as such, but there might be smaller block groups comprised of mostly “Green Acres” or “Salt of the Earth” which are lost in the larger enumeration unit. Figure 3.2 shows Guilford County color coded by the Community Tapestry.

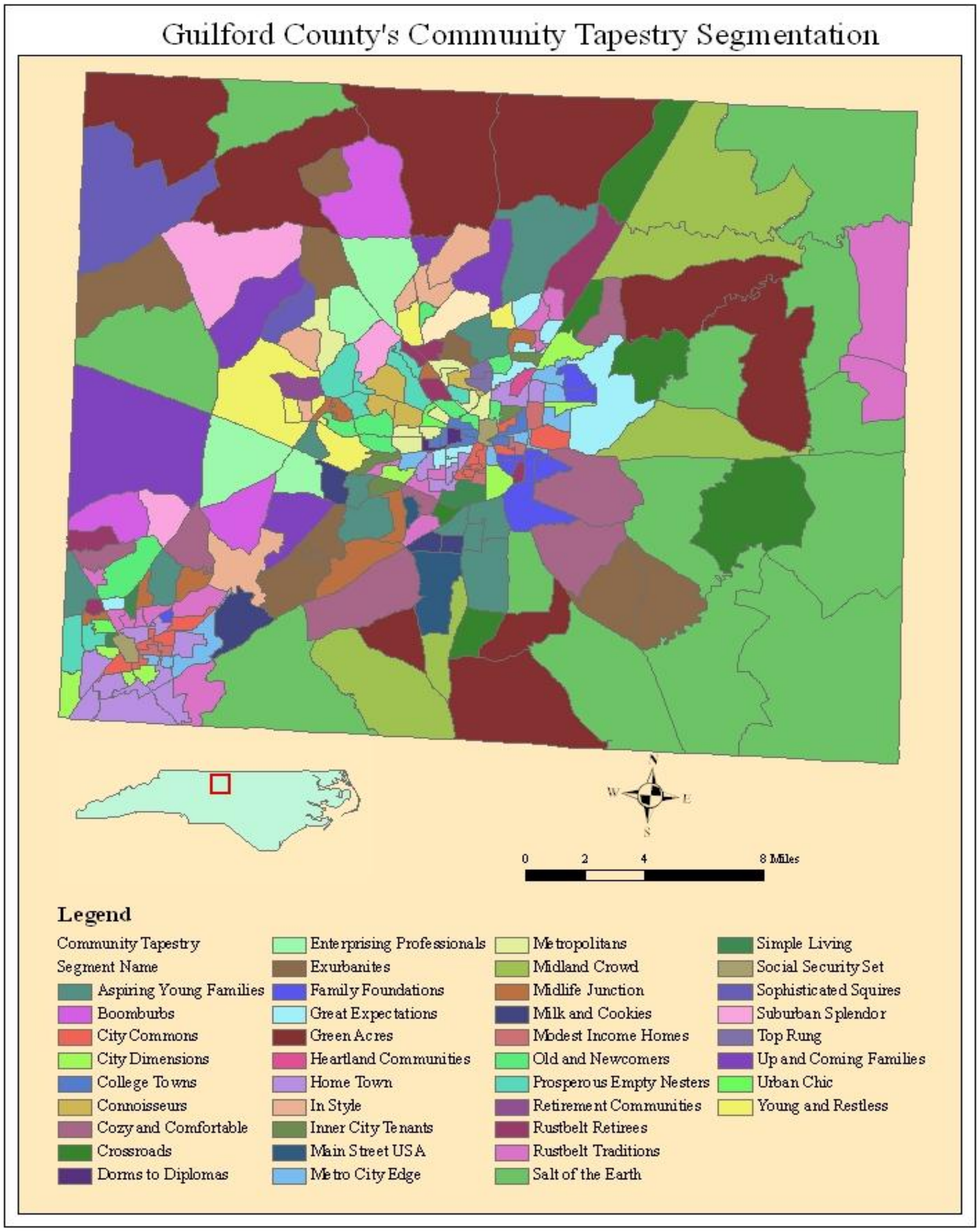

Figure 3.2. Guilford County’s Community Tapestry Segmentation

After the three trade areas were defined, they were combined using the union and dissolve tools in GIS. This created an overall trade area for the company's well-established locations. Because the data were analyzed at the block group level, the next step was to select the block groups intersecting the composite trade area. In all, 161 block groups in Guilford, Forsyth, and Rockingham Counties intersected the trade area. The block groups were clipped to the shape of the extent of the trade area (Figure 3.3). Because the trade area bisects many block groups, the number of households attributed to those block groups does not always accurately reflect the number of households in the trade area for that block group.

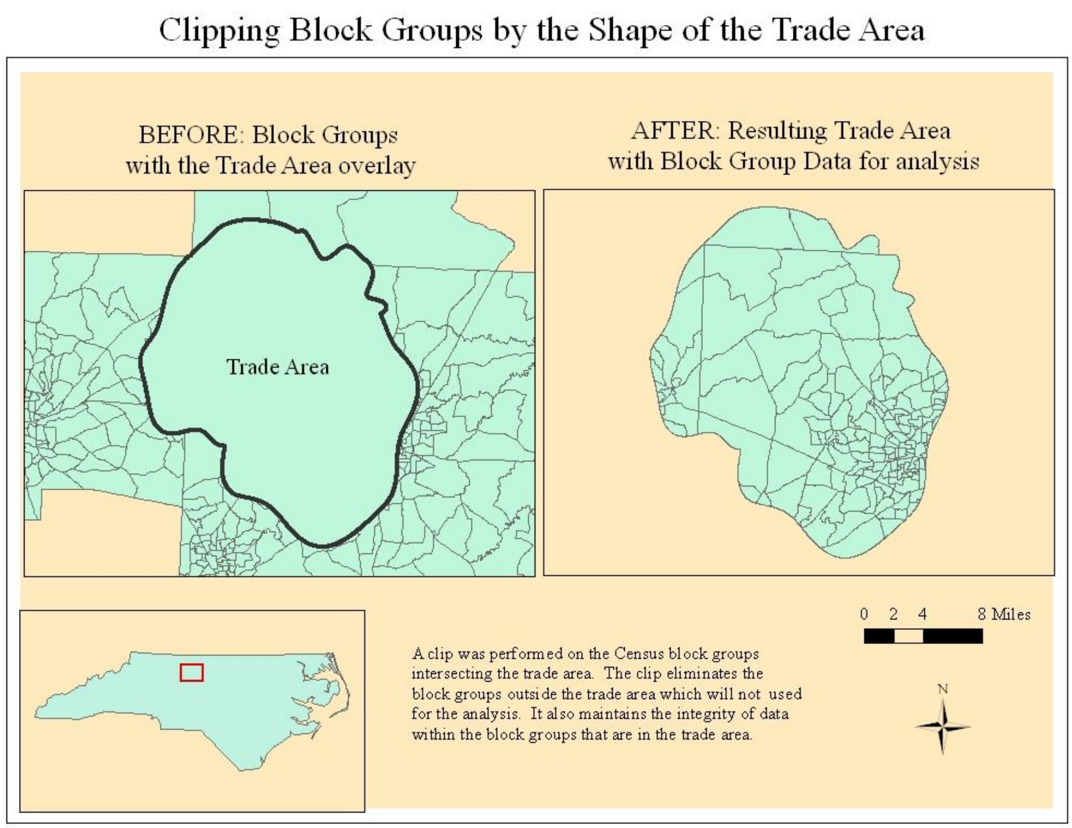

Figure 3.3. Clipping Block Groups by the Shape of the Trade Area

Instead of estimating this number, the census blocks were placed inside the block groups in the shape of the edited trade area. Each block has the number of households in its attribute data. The attributes of the blocks completely contained in the block group were spatially joined to the block groups and a field was created and named "Total Households." Bisected blocks required estimation, which was computed by calculating the percentage of the area within the trade area and multiplying this ratio by population within each block. The computed number of households was added to the "Total Households" field and represented an accurate estimation of households within each bisected block group.

The households from each block group completely contained in the trade area were added along with those households estimated from the bisected block groups producing 104,472 households in the trade area. The total number of customers for the three branches is 3,803, so their overall market penetration within their trade area is 3.6% (Equation 3.1).

$$\frac{\text{Customer Households}}{\text{Total Households}} = \frac{3{,}803}{104{,}472} = 3.6\%$$

Equation 3.1

It is important to note that the numbers of customers produced by Equation 3.1, (3,803) is not the same as the number produced by Applebaum's method in Table 3.1 (3,461). A trade area was made for each individual branch, but a branch's trade area may contain additional customers in the trade areas of the other branches. Because they were viewed as a composite trade area, all of the customers for the three branches were included. For example, Branch A had 1,788 customers in its own trade area, but also had customers in the trade areas for Branch B and Branch C. Of course, customers in areas where the trade areas overlapped should not be counted more than once. The same analogy applies to branches B and C. Equation 3.2 is the formula behind this logic.

$$\sum_{\text{Branch A}}^{\text{Branch C}} \mathrm{N} + \mathrm{M} - \mathrm{O}$$

Equation 3.2

Where: N = The Number of Customers for a branch within its own trade area.
M = The number of customers for a branch within the other branches' trade areas.
O = The number of customers contained in the overlap.

Some of the blocks that were bisected by the trade area were removed from the analysis because not enough of the area was contained within the trade area to be meaningful. Blocks with less than half of the total area inside the trade area were excluded and the households and customers in those blocks were not included. This created more precision in the analysis by eliminating numerous instances where household numbers were estimated. Only 29 customers were eliminated on the fringe of the trade area by removing these blocks. When the same process was applied to the block groups, an additional 35 customers were removed. Applebaum (1966b) suggested that 60-70% of customers are enough to define a primary trade area. From the 5,003 known addresses, 3,803 were used in the original trade area. After subtracting the 64 that were removed from the fringe, 3,739 remained, which is 74.7% of the matched addresses.

This amount is more than enough to determine the primary trade area, so removing 64 customers did not impact the integrity of the primary trade area.

With the elimination of blocks and block groups that were not at least 50% covered by the trade area, the attribute data were ready to be manipulated for analysis. The number of households had already been calculated for each block group. The total number of trade area customers was also known. The customers were then spatially joined to each block group creating a "Customer Count" field. A "Penetration Ratio" field was created for each block group. Finally, the Community Tapestry Data was joined to the block groups. Five fields were exported for the analysis: block group identification numbers, market penetration ratio, dominant tapestry codes, life mode codes, and urbanization codes.

At this point, the data were prepared for analysis; however, one final step was required. After the trade area was created, it needed editing to ensure it did not include areas where people cannot live. For example, in the trade area of Branch A, Piedmont Triad International Airport, Lake Higgins and Reedy Fork Creek are examples of such locations. The boundaries of the trade area were edited to exclude these topographic barriers. Shaping the trade area to exclude human and topographic barriers was performed after the penetration ratio was calculated because exclusion prior to the calculation would introduce unnecessary error. Figure 3.4 shows the final shape of the trade area.

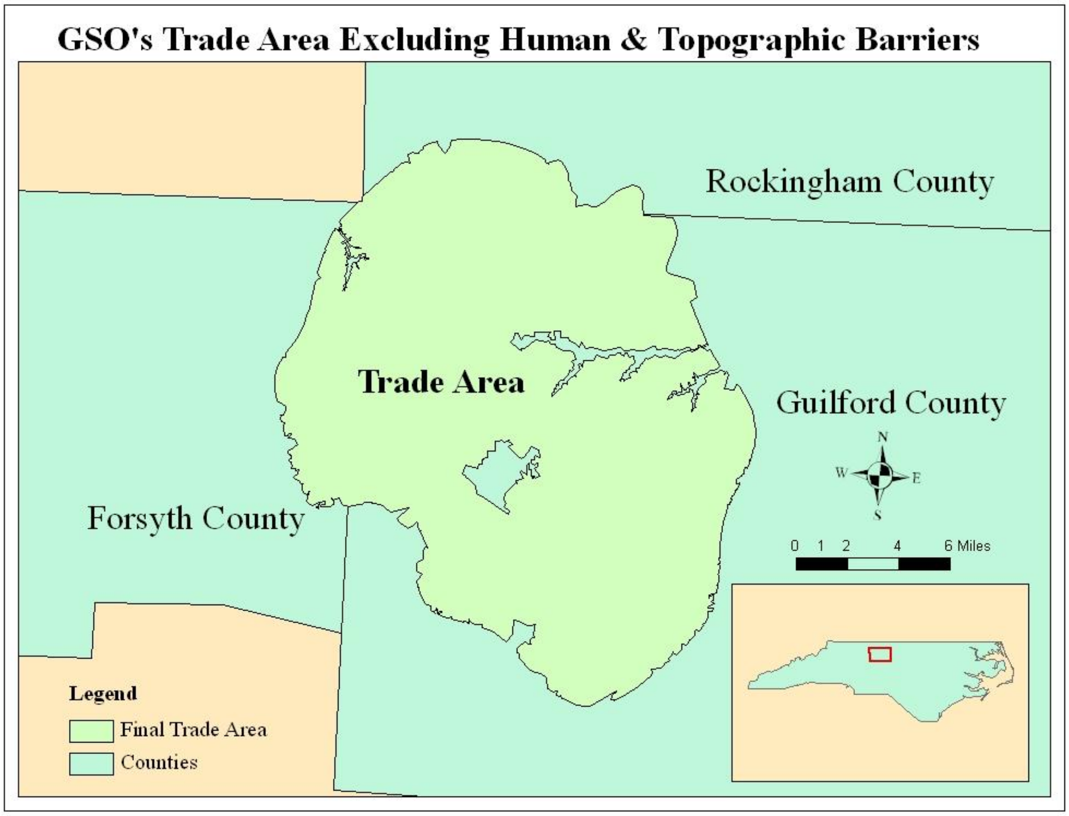

Figure 3.4. GSO’s Trade Area Excluding Human and Topographic Barriers

CHAPTER IV

ANALYSES & RESULTS

A univariate Analysis of Variance (ANOVA) was performed on market penetration ratios using the Community Tapestry codes as the independent variable. A second ANOVA was also performed on the market penetration ratios using the Life Mode Codes and Urban Codes as the independent variables. A final ANOVA categorized the Urban and Life Mode Codes into four groups based on location and affluence. The following describes, in detail, the variables and the three analyses used in this study.

4.1 Description of the Variables

4.1.1 Community Tapestry Codes

The Community Tapestry Codes represent 65 unique market segments in the Community Tapestry lifestyle/life-mode dataset (Appendix A). Each segment was derived from over 60 different variables, such as age, income, education, family size, etc. The segments were made using cluster analysis as discussed in Chapter 2, Section 6. These variables were believed to be determinants of consumer behavior, so the most fundamental task of the Community Tapestry was to group people into consumer behavior types. The Community Tapestry dataset provided the independent variables that were used as main effects in all three analyses.

The Community Tapestry Code contains Life Mode and Urban Codes by block groups. These codes represent summary groupings of the 65 Tapestry segments. The 12 Life-Mode groups are based on lifestyle and life stage and provide a broader view of the Tapestry. Segments in the same Life-Mode summary group share traits such as age or affluence. The 11 Urban Codes are based on geographic location, physical features and income. Tapestry segments in the same Urbanization summary group share a locale; however, the 1 or 2 appearing after the summary group name indicated the relative affluence of the group with 1 being more affluent than 2 (ESRI 2007).

4.1.2 Market Penetration Ratios

The market penetration ratios provided the dependent variable for all three analyses. In each block group, the penetration ratio is defined as the number of unique household addresses at which a GSO customer resides divided by total households. See Chapter 3 for details. This ratio represents the best measure of GSO's current market share within each block group in its trade area. This ratio has been used successfully in a number of site selection studies (Applebaum 1966b, Jones and Simmons 1989).

4.2 Analysis of Variance - Community Tapestry

All 65 segments of the Community Tapestry were considered in the analysis to explain the variance in market penetration ratios for GSO's trade area. Given the extent of the trade area, 44 of the Tapestry segments were not included in the analysis. Section 4.3 will describe how relevant information can still be extracted for GSO about these segments using the existing market penetration data. The remaining 21 were used as the main effect LIFESTYLE (Table 4.1). A histogram indicated the market penetration ratios were skewed to the right (Figure 4.1). The skewness was 5.48 which indicated it was not normally distributed, and thus, a natural logarithmic transformation (ln) was applied. Figure 4.2 shows the new distribution with a relatively modest skewness of -0.26. The results of the transformation were used as the dependent variable and called LOGPENT.

Table 4.1. Main Effect LIFESTYLE

Main Effect LIFESTYLE: F= 6.54, F > P = <0.0001		
Tapestry Segment	N	Mean Penetration Ratio
Green Acres	4	21.30%
Exurbanites	6	17.51%
Prosperous Empty Nesters	6	3.03%
In Style	6	2.61%
Enterprising Professionals	4	1.79%
Up & Coming Families	4	2.74%
Connoisseurs	4	1.57%
Rustbelt Retirees	3	1.28%
Crossroads	3	1.20%
Hometown	5	0.64%
Metropolitans	9	1.52%
Midlife Junction	7	1.18%
Rustbelt Traditions	5	0.35%
Milk & Cookies	4	0.36%
Great Expectations	6	0.24%
Old & Newcomers	8	0.73%
Young & Restless	7	0.71%
Aspiring Young Families	9	0.62%
College Towns	6	0.24%
Inner City Tenants	4	0.19%
City Commons	5	0.17%

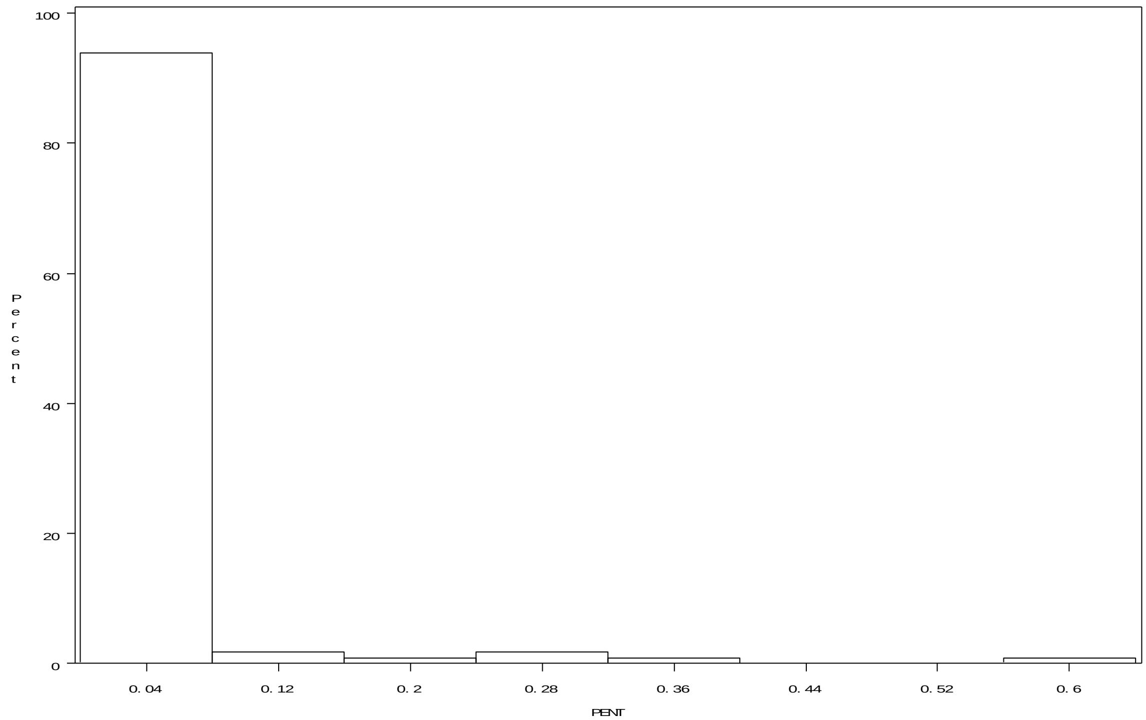

Figure 4.1. PENT

Eighteen of the original observations had a penetration ratio of zero because GSO had no clients in those 18 block groups (Appendix B). A log does not apply to values of zero, but they could not be excluded from the analysis. Zero is a valid value in this instance, so these were given a false value of 0.0002. They are visible on the far left of the distribution in Figure 4.2. The lowest nonzero market penetration ratio was 0.000892. Transforming this value created a new value of -7.02198, and the value of the log 0.0002 was -8.51719. These values were far enough apart to ensure the original values of zero are not interfering with the other values, but were close enough to create the relatively modest skewness of -0.26.

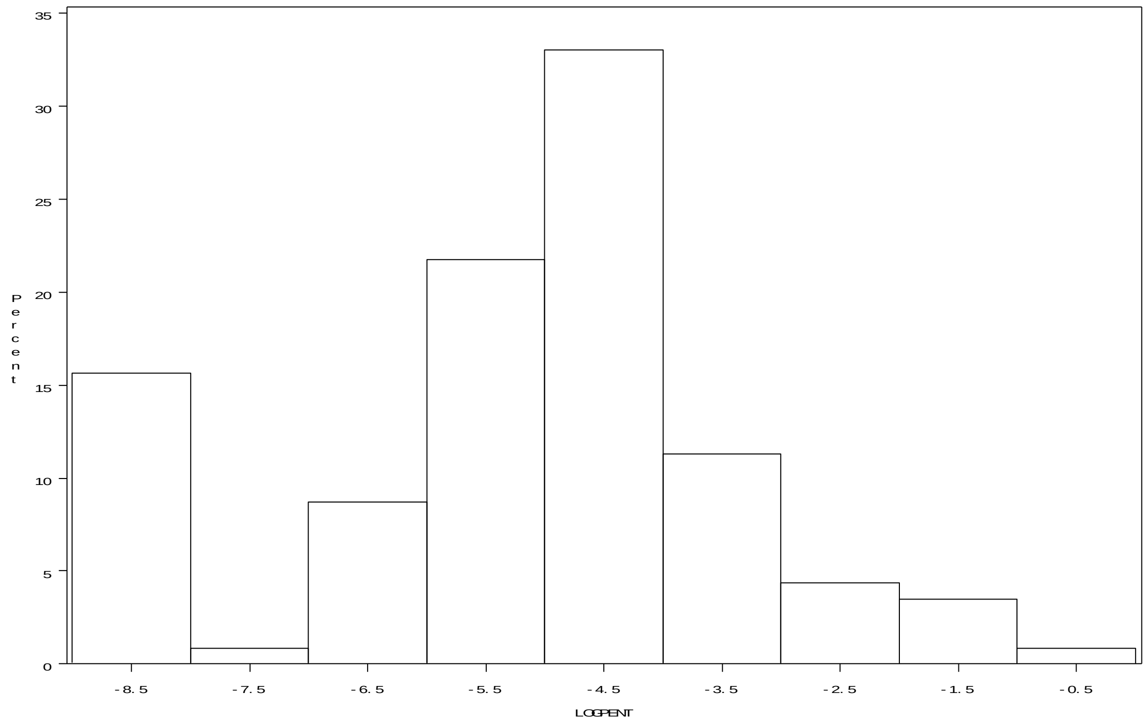

Figure 4.2. LOGPENT

Main effect LIFESTYLE was significant at $\alpha = 0.1$ (F = 6.54, P > F = <0.0001). The original penetration ratio means are shown in Figure 4.3. A post-hoc Least Significant Difference (LSD) multiple range test of means indicated that Green Acres and Exurbanites were in the significantly highest grouping. The LSD also showed seven tapestry segments to be in the significantly lower than the other segments. These were Young & Restless, Aspiring Young Families, Milk & Cookies, Rustbelt Traditions, College Towns, Hometown, Inner City Tenants, City Commons, & Great Expectations (Appendix C).

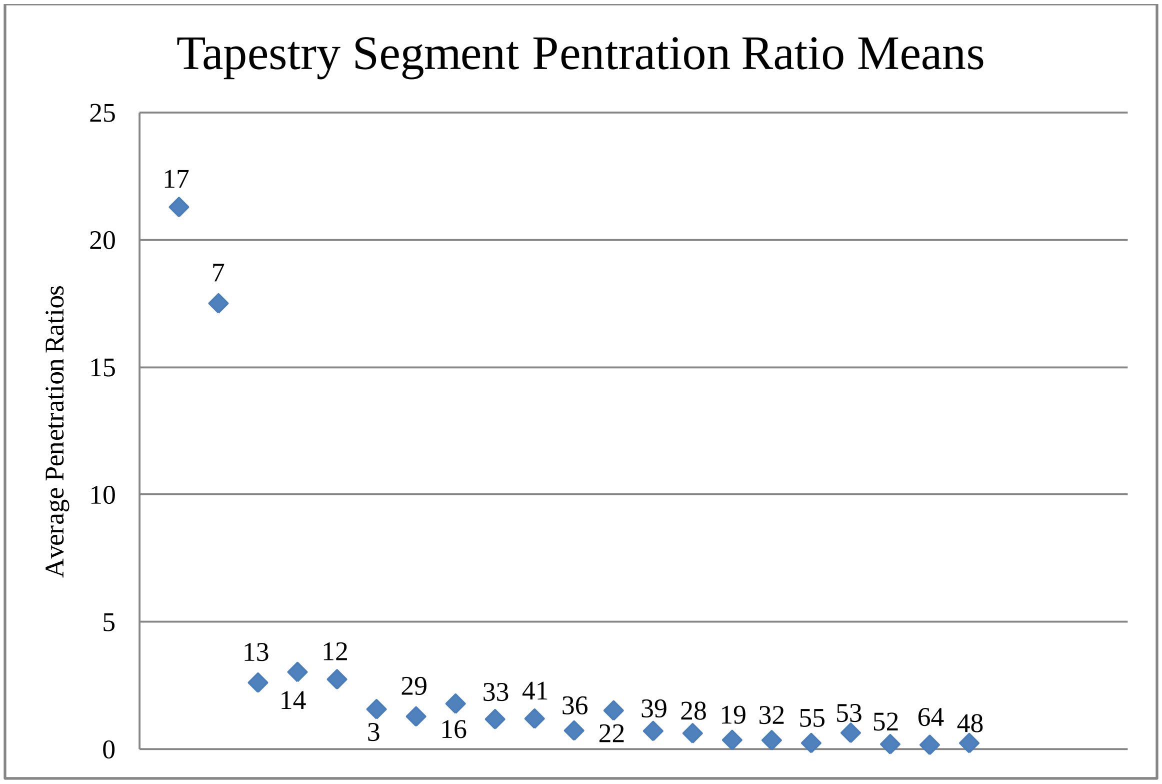

Figure 4.3. Tapestry Segment Penetration Ratio Means

These results seem reasonable because Green Acres (17) and Exurbanites (7) are affluent, educated segments with high employment rates. They also have median ages in the 40's meaning many are well-established families that have built savings and own homes. These are ideal clients for a small bank, such as GSO. The results of the lower groups such as Inner City Tenants (52) and Great Expectations (48) can be explained because these are less affluent, younger, with many renting from multifamily housing units. Most do not have college educations, and many are without a high school education.

4.3 Analysis of Variance – Life Mode & Urban Codes

The summary groups of Life Mode and Urban Codes were considered in this analysis to explain the variance in market penetration ratios for GSO's trade area. Given the extent of the trade area, all twelve Life Mode groups were used as the main effect LIFECODE (Table 4.2). Nine of the eleven Urban groups were used as the Main effect URBANCODE (Table 4.3). A histogram indicated the market penetration ratios were skewed to the right (Figure 4.4). The skewness was 4.03, which indicated it was not normally distributed, and thus, a base-ten logarithmic transformation was applied. Figure 4.5 shows the new distribution with a relatively modest skewness of -0.18. The results of the transformation were used as the dependent variable and called LOGCODE.

Table 4.2. Main Effect LIFECODE

F= 8.77, F > P = <0.0001		
Life Mode Segment	N	Mean Penetration Ratio
High Society	16	13.87%
Upscale Avenues	17	6.52%
Factories and Farms	7	5.98%
American Quilt	4	4.98%
Senior Styles	12	2.23%
Metropolis	12	1.25%
Family Portrait	13	1.01%
Traditional Living	13	0.77%
Solo Acts	15	0.72%
Scholars & Patriots	8	0.69%
High Hopes	15	0.47%
Global Roots	6	0.21%

Table 4.3. Main Effect URBANCODE

F= 6.54, F > P = <0.0001		
Life Mode Segment	N	Mean Penetration Ratio
Rural 1	13	20.03%
Suburban Periphery 1	7	8.70%
Urban Outskirts 1	12	2.77%
Metro Cities 1	17	1.36%
Small Towns	13	1.20%
Suburban Periphery 2	12	1.00%
Metro Cities 2	4	0.74%
Principal Urban Centers 2	16	0.45%
Urban Outskirts 2	15	0.29%

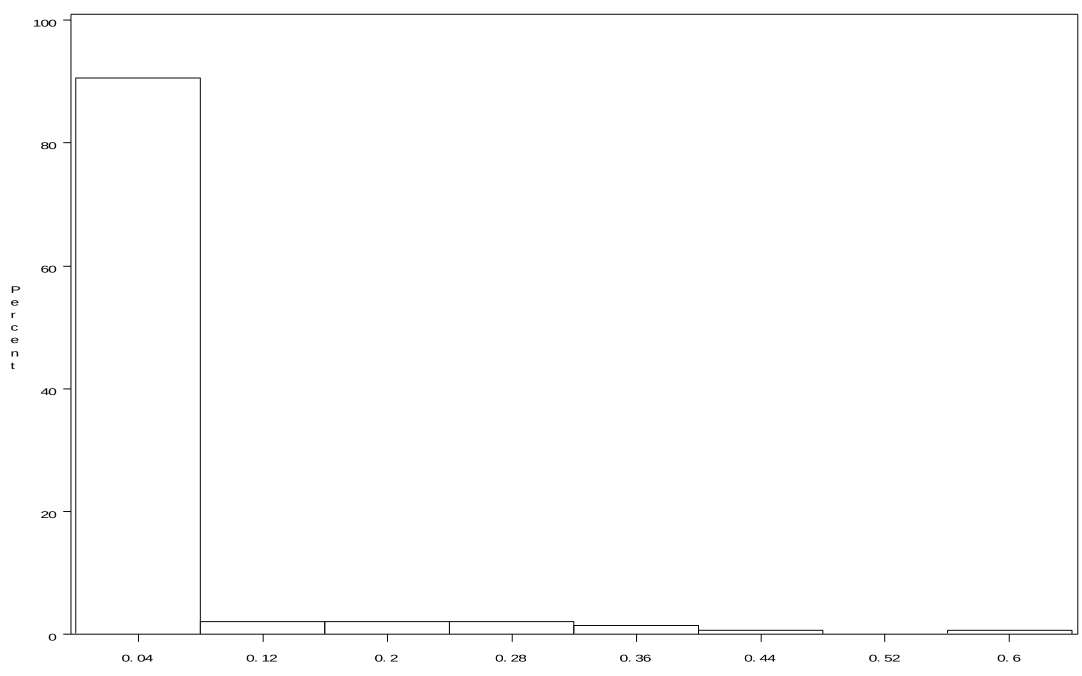

Figure 4.4. CODEPENT

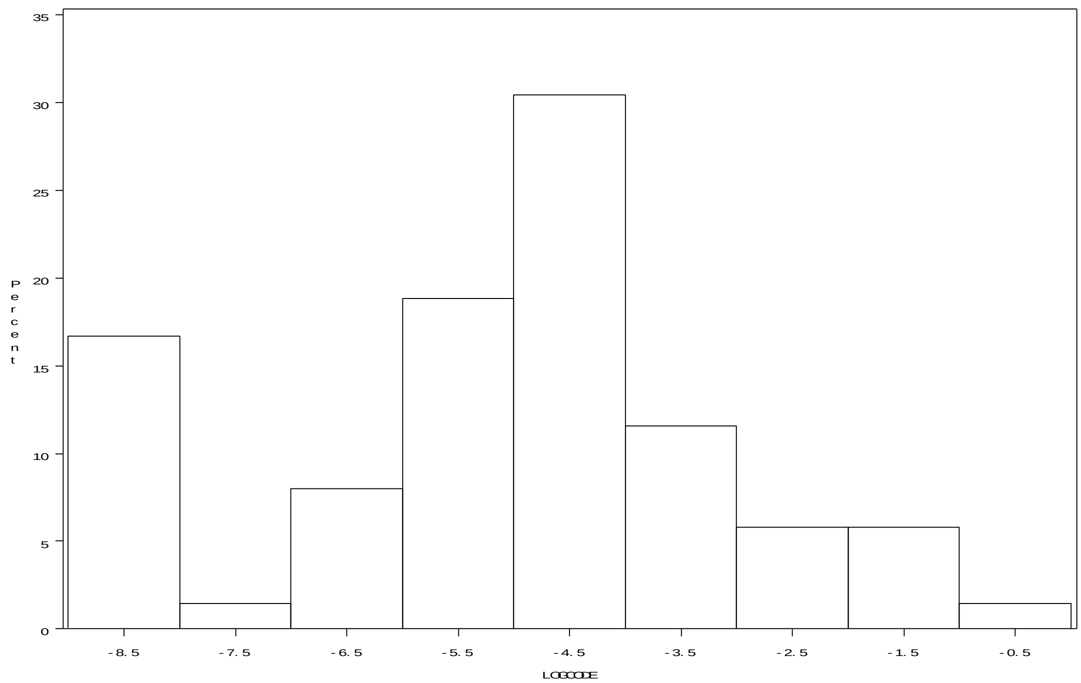

Figure 4.5. LOGCODE

Twenty-three observations had a penetration value of zero and were changed to 0.0002 because GSO had no clients in those 18 block groups (Appendix B). This was the same procedure as described in Section 4.2. The lowest nonzero penetration ratio did not change, so the values were far enough apart to ensure the original values of zero were not interfering with the other values, but are close enough to create the negligible skewness of -018.

Main effect LIFECODE was significant at $\alpha = 0.01$ ($F = 8.77$, $P > F = <0.0001$). The original penetration ratio means are shown in Figure 4.6. A post-hoc LSD multiple range test of means indicated that High Society, Upscale Avenues, and American Quilt were in the significantly highest grouping. These were LIFECODES 1, 2, and 12, respectively. The LSD also showed High Hopes, Scholars & Patriots, and Global Roots were among the Life Mode Groups with which GSO does not perform well (Appendix C).

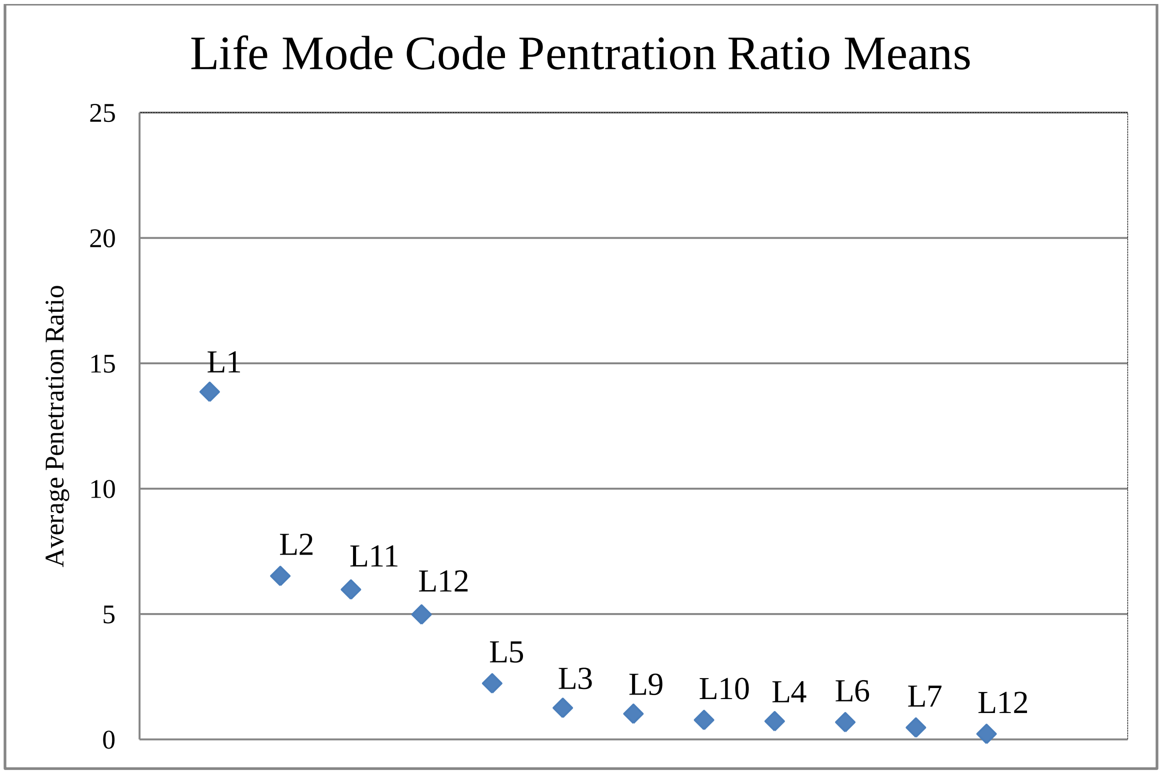

Figure 4.6. Life Mode Code Penetration Ratio Means

These results seem reasonable because High Society (L1) and Upscale Avenues (L2) are wealthy summary groups. They are well-educated with above average earnings. Both groups are mostly homeowners, with expensive homes. L1 has a median home value over $300,000. Investment is important in both groups and because both groups like to travel, both often save their earnings. For a mostly rural summary group, American Quilt (L12) is made up of older, relatively affluent, homeowners. Scholars & Patriots (L6), Global Roots (L8), and High Hopes (L7) are not life-mode summary groups with which GSO performs well because they have lower family incomes and most rent their dwellings. Many are not educated. Scholars and Patriots are marked by groups of students or military personnel who may not have the ability to own their home.

Main effect URBANCODE was significant at $\alpha = 0.01$ ($F = 7.19$, $P > F = <0.0001$). The original penetration ratio means are shown in Figure 4.7. A post-hoc LSD multiple range test of means indicated that Rural 1 and Suburban Periphery 1 were the Urban Codes with significantly higher penetration ratios than the other Urban Codes present in GSO's trade area, such as URBANCODES 7 and 10. The LSD also showed Codes representing more urban areas, such as Principal Urban Centers 2, Urban Outskirts 1 & 2, and Metro Cities 2 to be statistically significantly lower than the other groups (Appendix C).

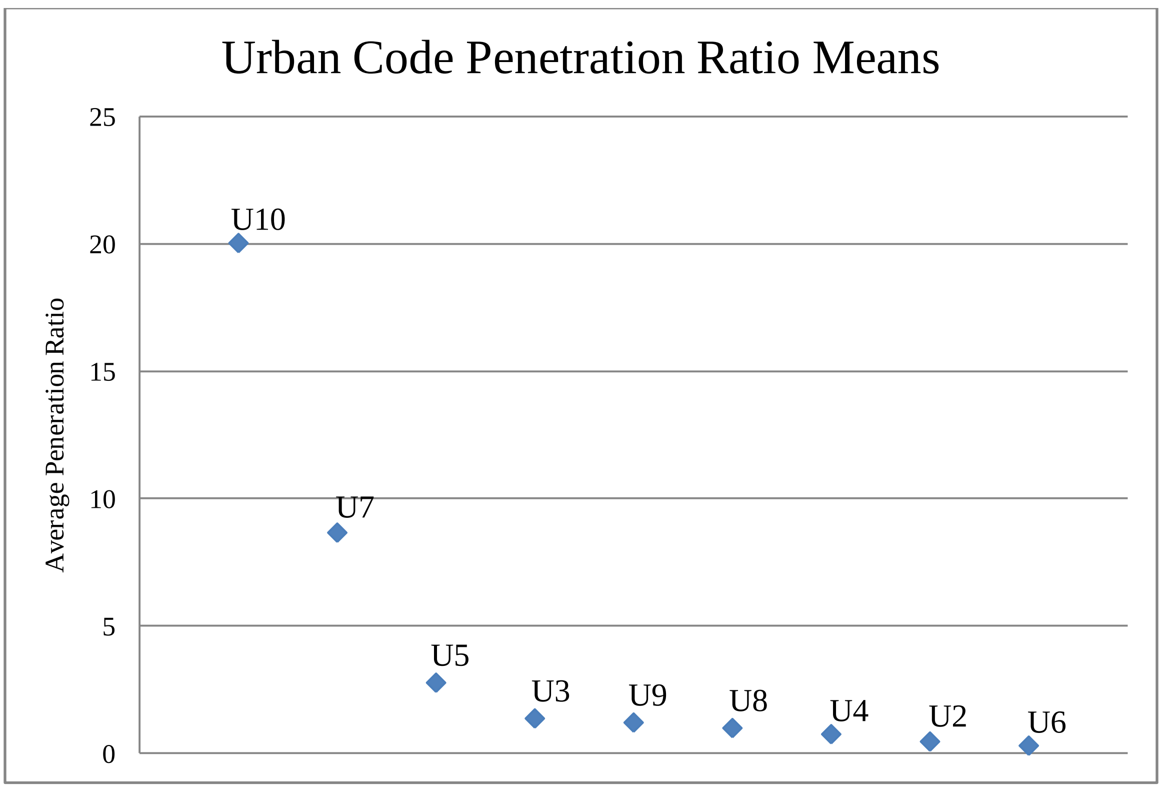

Figure 4.7, Urban Code Penetration Ratio Means

These results seem reasonable because Suburban Periphery (U7) and Rural 1 (U10) are almost entirely comprised of single-family housing units where inhabitants own their home. Most have a mortgage, and because many perform renovation and remodeling projects, many refinance and have second mortgages. Suburban Periphery 1 households are likely to invest and have financial planners. Many Rural 1 households are self-employed and deposit business revenues with a bank. The urban groups with which GSO does not perform well (U2, U4, U5 and U6) are made up of mostly multi-unit housing. Many households in these categories are renters. These are all also younger categories, and U2, U4 and U6 have median household incomes well below the national average.

4.4 Analysis of Variance – Addressing a Weakness

The analysis is not without weakness. Although the results are promising, the number of observations for all of the LIFESTYLE groups, and most of the LIFECODE, and URBANCODE groups were really too small to be statistically conclusive. The study is limited by the spatial extent of the trade area, so it would be beneficial to have a larger area, or several separate areas from which to perform the same analysis. The largest N for LIFESTYLE was nine. As mentioned previously, some segments were excluded from the analysis for only having an N of one or two occurrences in the trade area. Many Tapestry segments were not included at all.

An important benefit of the Summary groups was that they combined the Tapestry segments which increased the sample sizes. The largest N for LIFECODE was 17, but the smallest was four. For URBANCODE, the largest was 33 and the smallest was four

(Appendices A and C). These smaller Ns result from the inclusion of LIFESTYLE variables that were excluded from the original ANOVA. It is important to note that the LIFESTYLE groupings did not conflict with the groups made from LIFECODE and URBANCODE. For example, Green Acres and Exurbanites were the two highest LIFESTYLE groups. Green Acres' summary groups are Upscale Avenues and Rural 1, which also were statistically significantly higher in their respective tests. The Exurbanites group belongs to the summary groups High Society and Suburban Periphery 1, which were also statistically higher.

To check for the validity of the segmentation, the groups were again broadened for a final ANOVA. The segments were categorized by their URBANCODE, but the codes were combined from 12 groups to four (Table 4.4). The categories took into account both the urban and affluence aspect of the urban summary groups. The more urban groups were set aside from the more rural groups, and then those two sections were subdivided with the more affluent groups being separated from the less affluent. The four groups had much larger Ns and were used as the main effect UACODE.

Main effect UACODE was significant at $\alpha = 0.01$ ($F = 23.24$, $P > F = <0.0001$). A post-hoc LSD multiple range test indicated that Less Urban, More Affluent Tapestry Segments are significantly higher than the other Segments. The LSD showed there was a difference between the less affluent, rural segments and the less affluent, urban segments. There was no significant difference between the urban segments regardless of affluence.

Table 4.4. Main Effect UACODE

F = 23.24, P > F = <0.0001

Name	LSD	Mean	Total N	Description	Includes(N)
Rich Rural	A	-3.0328	33	Less Urban, More Affluent	Suburban Periphery 1(26), Rural 1(7)
Poor Rural	B	-5.2853	20	Less Urban, Less Affluent	Suburban Periphery 2(17), Rural 2(0), Small Towns (3)
Rich Urban	B	-5.5845	36	More Urban, More Affluent	Principal Urban Centers 1(0), Metro Cities 1(22), Urban Outskirts 1(14)
Poor Urban	B	-6.0048	49	More Urban, Less Affluent	Principal Urban Centers 2(6), Metro Cities 2(33), Urban Outskirts 2(10)

CHAPTER V

DISCUSSION AND CONCLUSIONS

5.1 Discussion

The objective of analog modeling is to find similarities between markets in the site selection process. The Community Tapestry provides a simple way of identifying these similarities. An advantage to the Tapestry is that it combines several useful pieces of data that could be used in the site selection process, such as median age, median household income, family size, etc. A disadvantage is that the specificity of those categories is lost, and depending on how important one of them may be to a firm, it might be too valuable to be combined with other characteristics.

One very important advantage in using the Community Tapestry as opposed to Census data involves the fact that the Tapestry identified groups which may not fit GSO's marketing profile. Traditional analog methods simply look for similarities between an existing trade area and the market around a potential site. Using Census data would make it less likely to find a demographic statistic a firm does not attract. It would not be prudent to find a Block Group with a low penetration ratio and deduce that the bank does not perform well with people of that block group's age, income, etc., because one block group is not an adequate sample size. Also, if all block groups with low ratios were excluded, it is likely there would be demographic differences among the block groups. If all of the demographics were assumed be negatively correlated to the GSO's success, then the list of demographics to avoid would likely become very large.

The Tapestry categorizes Block Groups with many similar characteristics to create larger sample sizes from which reasonable conclusions can be made. In this case, several prototypical family types were shown to be unprofitable. This allows GSO to see weaknesses in their own trade area and provides an opportunity for them avoid those weaknesses in a new market.

5.2 Conclusions for GSO

The Life Mode codes and Urban Codes were analyzed within the trade area of the bank in question, GSO, for two reasons. First, they might provide insight into why some tapestry segments perform better or worse than others. Second, this analysis allowed the inclusion of all block groups observed in the trade area. If a tapestry segment was excluded for not having a large enough sample from which to study, it was included in its summary groups with other segments that share its urban and life mode characteristics.

Categorizing each tapestry segment according to the results of the LSD multiple range tests of means allows these segments to be categorized (Table 5.2). The tapestry segments that were only included in the analysis of the summary groups can also be ranked (Table 5.3). The key to these categorizations is provided in Table 5.1.

Table 5.1. Key

Key	
	Statistically higher in their respective test.
	Above average, but not in highest group.
	Statistically lower in their respective test.

Table 5.2. Segments with at least Three Observations

Tapestry Segment	Life Mode Code	Urban Code
Green Acres	2	10
Exurbanites	1	7
In Style	2	7
Prosperous Empty Nesters	5	7
Connoisseurs	1	3
Up & Coming Families	9	7
Enterprising Professionals	2	3
Crossroads	12	9
Rustbelt Retirees	5	3
Midlife Junction	10	8
Old & Newcomers	4	4
Metropolitans	3	8
Young & Restless	4	4
Milk & Cookies	9	3
Hometown	11	8
Aspiring Young Families	7	4
Inner City Tenants	8	4
Rustbelt Traditions	10	5
City Commons	9	2
College Towns	6	6
Great Expectations	7	5

Dominant tapestry codes are more specific in explaining the variance of the penetration ratios, but this does not mean the Life Mode and Urban Codes are not valuable. For example, the tapestry segment Midland Crowd was not included in the original ANOVA using the tapestry codes because there were only two block groups in the trade area with that designation. Midland Crowd's summary group, American Quilt was in the significantly higher group when Life Mode Codes were analyzed. Its Urban Code, Rural 1, was in the highest grouping. This suggests Midland Crowd is likely as important to GSO as Exurbanites or Green Acres which were in the original analysis.

Table 5.3. Segments with less than Three Observations

Tapestry Segment	Life Code	Urban Code
Midland Crowd	12	10
Suburban Splendor	1	7
Sophisticated Squires	1	7
Salt of the Earth	11	10
Urban Chic	2	3
Cozy & Comfortable	2	8
Boomburbs	1	5
Retirement Communities	5	4
Southern Satellites	11	11
Social Security Set	5	2
City Dimensions	8	4
Dorms to Diplomas	6	4
Main Street	10	5
Metro Cities	3	6
Simple Living	5	6
Modest Income Homes	3	6

Knowing the results of the Life Mode Code and Urbanization Code analysis allows all tapestry segments to be ranked regardless of their number of observations, or even their inclusion in the trade area. Tapestry segments that are not well represented in GSO's trade area might be more common in other markets into which GSO is interested in expanding. It would be unwise to ignore segments that were not used in the first analysis if information can be obtained about them from their Life Mode and Urbanization Codes. Once GSO knows their critical tapestry segments, life modes, and urban locales, they can rank each tapestry segment and identify customers that would best fit their ideal customer profile.

Any ranking of tapestry segments will be subjective, but when using analogs, subjectivity is required (Applebaum 1966b). A subjective ranking and grouping of all tapestry segments in the trade area is provided in Table 5.4. Group One contains segments with which GSO performs best. In the future, GSO should make it a priority to locate among or near these people. GSO would likely be successful with groups Two and Three if they were to locate closer to these segments. Groups Six through Eight are segments with which GSO does not perform well.

These analyses were not entirely representative of the community tapestry. One Urban summary group, Principal Urban Centers 1, was not represented in the trade area, so potential among tapestry segments characterized by that code is still largely unknown. Also, there was one block group of the segment called Southern Satellites. It was the only block group in the trade area with the Urban Code Rural 2, so that summary group was not included in the Urban Code analysis. Table 5.4 lists Southern Satellites in Group Four, but if a larger sample was known, the Urban Code could be coded differently, changing its location in Table 5.4 and making it a priority for consideration in a new market.

There is enough information to rank and group all but ten segments, even if they are not located in the banks' trade area. This is useful information. If the bank moves into new markets, there is no guarantee it will only have the same segmentation as the study area. Several of these segments have Urban Codes designating them as Principal Urban Centers 1, which are not present in any of the potential markets GSO has considered. GSO is more likely to encounter Rural 2 as it expands into new markets.

Table 5.4. Ranking of All Segments in GSO Trade Area

Group	Tapestry Segment	Life Code	Urban Code
1	Green Acres	2	10
	Exurbanites	1	7
	Midland Crowd	12	10
2	In Style	2	7
	Suburban Splendor	1	7
	Sophisticated Squires	1	7
3	Prosperous Empty Nesters	5	7
	Connoisseurs	1	3
	Up & Coming Families	9	7
	Salt of the Earth	11	10
4	Enterprising Professionals	2	3
	Crossroads	12	9
	Urban Chic	2	3
	Cozy & Comfortable	2	8
	Rustbelt Retirees	5	3
	Retirement Communities	5	4
5	Boomburbs	1	5
	Social Security Set	5	2
	Southern Satellites	11	11
	Midlife Junction	10	8
	Old & Newcomers	4	4
	Metropolitans	3	8
	Young & Restless	4	4
6	Milk & Cookies	9	3
	Hometown	11	8
	City Dimensions	8	4
	Dorms to Diplomas	6	4
	Aspiring Young Families	7	4
	Inner City Tenants	8	4
7	Main Street	10	5
	Metro Cities	3	6
	Simple Living	5	6
	Modest Income Homes	3	6
8	Rustbelt Traditions	10	5
	City Commons	9	2
	College Towns	6	6
	Great Expectations	7	5

5.3 Future Research Questions

Would the methodology produce similar results if the data sources were different? The analysis has proven the Community Tapestry is a valuable tool in determining potential customers that result in higher profits for GSO. If a retail chain used its customer data and the Community Tapestry, ANOVA might produce statistically significant groupings for them as well. Also, if GSO used a different lifestyle segmentation dataset an ANOVA performed on those segments should produce similar results to those of this study.

Another important consideration is the effect distance has on market penetration. Does market penetration of different tapestry segments decay at different rates? This might be influenced by the fact that penetration ratios are highest in Block Groups adjacent to the branches. Penetration ratios are highest in block groups next to the branches, and generally lowest in Block Groups on the periphery of the trade area. For a block group next to the branch, answering the distance question could determine whether the Tapestry determines GSO's success or if the bank is successful in those block groups only because they are adjacent to a branch.

Can a trade area actually be predicted using the Community Tapestry? If, for example, Green Acres' mean penetration and distance decay rate are known, it is quite possible that points representing potential customers could be randomly distributed within a Green Acres block group near a proposed site based on its distance from the site, and its population base. If this were done for each block group around the site, a firm might be able to delineate an estimation what its trade area might look like in the new market.

Lastly, as GSO's branches that were excluded from the analysis develop, would they create higher rates of market penetration in more urban areas? Much of GSO's success is with more affluent, less urban households. They might they become successful with more affluent, urban households as well and this could be reflected with improved market penetration in different tapestry segments and summary groups. Because there are similarities between the sites GSO is considering locating new branches and the markets serviced by their excluded branches, this might be valuable information for their management team and board.

REFERENCES

Applebaum, William and Saul B. Cohen. 1960. Evaluating Store Sites and Determining Store Rents. Economic Geography, 36 (1): 1-35.

Applebaum, William. 1954. 'Marketing Geography.' American Geography: Inventory and Prospect, 245-251.

Applebaum, William. 1965. Can store location research be a science? Economic Geography, 41 (3): 234-237.

Applebaum, William. 1966a. Guidelines for a Store-Location Strategy Study. The Journal of Marketing, 30 (4): 42-45.

Applebaum, William. 1966b. Methods for Determining Store Trade Areas, Market Penetration, and Potential Sales. Journal of Marketing Research, 3 (2): 127-141.

Black, William C, Lyman E. Estlund & Robert & Westbrook. 1985. Spatial Demand Models in an Intrabrand Context, The Journal of Marketing, 49 (3): 106-113.

Blair, Edward. 1983. Sampling issues in Trade Area Maps Drawn from Shopper Surveys. The Journal of Marketing, 47 (1): 98-106.

Berry, Brian J.L 1967. Geography of Market Centers and Retail Distribution. Englewood Cliffs, NJ: Prentice Hall. 1967.

Boufounou, P.V. 1995. Evaluating Bank Branch Location and Performance: A Case Study. European Journal of Operational Research. 87: 389-402.

Breheny, Michael J. 1998. Practical Methods of retail location analysis: a review. In Store Choice, Store Location, and Market Analysis, ed. Neil Wrigley, 39-86. New York: Routledge.

Brown, Stephen. 1992. Retail Location: A Micro-Scale Perspective. Brookfield, VT: Ashgate Publishing.

Converse, P.D. 1943. A Study of Retail Trade Areas in East Central Illinois. University of Illinois, Bureau of Economic and Business Research, Business Studies, 2.

Engle, Janet. 2007. How to Open an Operate a Financially Successful Retail Business. Ocala, FL: Atlantic Publishing Group.

Fetto, J. 2004. American Neighborhoods' First Page. American Demographics, 26: 32.

Fotheringham, A. Stewart. 1998. Market share analysis techniques: a review and illustration of current U.S. practice. In Store Choice, Store Location, and Market Analysis, ed. Neil Wrigley, 120-159. New York: Routledge.

Freehling, John. 1993. Use Drive Times to Build Trading Areas and Market Segments. Marketing News, 27(10). 2.

William J. Reilly, Methods for Study of Retail Relationships (Austin: The University of Texas, Bureau of Business Research, Research Monograph, No. 4, 1929).

Goss, Jon. 1995. 'We Know Who You Are and We Know Where You Live:' The Instrumental Rationality of Geodemographic Systems. Economic Geography, 71 (2) 171-198.

Haines, George H., Leonard S. Simon, Marcus Alexis. 1972. Maximum Likelihood Estimation of Central-City Food Trading Areas. Journal of Marketing Research, 9 (2): 154-159.

Huff, David L. 1963. A probabilistic analysis of shopping trade areas. Land Economics. 39: 81-90.

Huff, David L. 1964. Estimating a Trading Area. The Journal of Marketing, 28 (3): 34-38.

Jones, K.G., and Mock, D.R. 1984. Evaluating Retail Trading Performances. In Store Location and Store Assessment Research, eds. R.L. Davies and D.S. Rogers. Store Location and Store Assessment Research, 333-360. New York: John Wiley and Sons.

Jones, Ken and Jim Simmons. 1987. Location, Location, Location: Analyzing the Retail Environment. New York: Methuen.

Min, Hokey. 1989. A Model-Based Decision Support System for Locating Banks. Information and Management, 17 (4): 207-215.

Kivell, P.T. and G. Shaw. 1984. The Study of Retail Location. In Retail Geography, ed. John A. Dawson, 95-155. New York: Halsted Press.

Kolter, P. 1971. Marketing Decision Making: A Model Building Approach. New York: Holt, Rinehart and Winston.

Lee, Ming Long and R. Kelley Pace. 2005. Spatial Distribution of Retail Sales. The Journal of Real Estate Finance and Economics, 31(1): 53-69.

Munroe, Keith. 2004. Meaningful Metrics: Site Evaluation and Market Optimization Models. In The Power of Place: Advanced Customer and Location Analytics for Market Planning, 127-156. San Diego: Claritas.

Peterson, Keith and Scott Munroe. 2004. Macro to Micro: Customer-Driven Retail Growth Planning. In The Power of Place: Advanced Customer and Location Analytics for Market Planning, 9-24. San Diego: Claritas.

Scott, Peter. 1970. Geography and Retailing. Chicago: Aldine Publishing.

Smith W. 1956. Product Differentiation and Market Segmentation as Alternative Marketing Strategies. Journal of Marketing, 21: 3-8.

Soenen, Luc A. 1974. Locating Bank Branches. Industrial Marketing Management. 3. 211-227.

Stanley, T. J., and M. A. Sewall. 1976. Image Inputs to a Probabilistic Model: Predicting Retail Potential. Journal of Marketing, 40, 48-53.

Stahel, P. (1969, March). Finding the Best Bank Site. Banking, 61(9), 54.

Stevens, B.H. 1985. Location of Economic Activities: the JRS contribution to the research literature. Journal of Regional Science, 25: 663-685.

Tedrow. David. 2004. Customer Profiling: Collecting and analyzing the right data. In The Power of Place: Advanced Customer and Location Analytics for Market Planning, 39-56. San Diego: Claritas.

Thompson, John S. 1982. Site Selection. New York: Lebhar-Friedman Books.

Vyncke, Patrick. 2002. Lifestyle Segmentation: From Attitudes, Interests and Opinions to Values, Aesthetic Styles, Life Visions and Media Preferences. European Journal of Communication, 17 (4): 445-463.

Appendix A

Community Tapestry Lifestyle Market Segments

Segment	Tapestry Segment Name	Life Mode Name	Urban Name	Median Age	Income
1	Top Rung	(L1) High Society	(U3) Metro Cities 1	42.4	High
2	Suburban Splendor	(L1) High Society	(U7) Suburban Periphery 1	41.5	High
3	Connoisseurs	(L1) High Society	(U3) Metro Cities 1	47.3	High
4	Boomburbs	(L1) High Society	(U5) Urban Outskirts 1	33.7	Upper Middle
5	Wealthy Seaboard Suburbs	(L1) High Society	(U5) Urban Outskirts 1	42.3	Upper Middle
6	Sophisticated Squires	(L1) High Society	(U3) Metro Cities 1	38.3	Upper Middle
7	Exurbanites	(L1) High Society	(U7) Suburban Periphery 1	44.8	Upper Middle
8	Laptops and Lattes	(L4) Solo Acts	(U1) Principal Urban Centers 1	38.5	Upper Middle
9	Urban Chic	(L2) Upscale Avenues	(U3) Metro Cities 1	42.0	Upper Middle
10	Pleasant-Ville	(L2) Upscale Avenues	(U3) Metro Cities 1	39.8	Upper Middle
11	Pacific Heights	(L2) Upscale Avenues	(U1) Principal Urban Centers 1	39.1	Upper Middle
12	Up and Coming Families	(L9) Family Portrait	(U7) Suburban Periphery 1	31.9	Upper Middle
13	In Style	(L2) Upscale Avenues	(U7) Suburban Periphery 1	40.3	Upper Middle
14	Prosperous Empty Nesters	(L5) Senior Styles	(U7) Suburban Periphery 1	48.7	Upper Middle
15	Silver and Gold	(L5) Senior Styles	(U7) Suburban Periphery 1	59.7	Upper Middle
16	Enterprising Professionals	(L2) Upscale Avenues	(U3) Metro Cities 1	32.4	Upper Middle
17	Green Acres	(L2) Upscale Avenues	(U10) Rural 1	40.7	Upper Middle

18	Cozy and Comfortable	(L2) Upscale Avenues	(U8) Suburban Periphery 2	42.1	Upper Middle
19	Milk and Cookies	(L9) Family Portrait	(U3) Metro Cities 1	34.0	Middle
20	City Lights	(L3) Metropolis	(U1) Principal Urban Centers 1	37.8	Middle
21	Urban Villages	(L9) Family Portrait	(U1) Principal Urban Centers 1	30.5	Middle
22	Metropolitans	(L3) Metropolis	(U3) Metro Cities 1	37.7	Middle
23	Trendsetters	(L4) Solo Acts	(U1) Principal Urban Centers 1	35.5	Middle
24	Main Street, USA	(L10) Traditional Living	(U5) Urban Outskirts 1	36.8	Middle
25	Salt of the Earth	(L11) Factories and Farms	(U10) Rural 1	41.1	Middle
26	Midland Crowd	(L12) American Quilt	(U10) Rural 1	37.0	Middle
27	Metro Renters	(L4) Solo Acts	(U1) Principal Urban Centers 1	33.8	Middle
28	Aspiring Young Families	(L7) High Hopes	(U4) Metro Cities 2	30.6	Middle
29	Rustbelt Retirees	(L5) Senior Styles	(U8) Suburban Periphery 2	45.0	Middle
30	Retirement Communities	(L5) Senior Styles	(U4) Metro Cities 2	51.9	Middle
31	Rural Resort Dwellers	(L12) American Quilt	(U10) Rural 1	47.1	Middle
32	Rustbelt Traditions	(L10) Traditional Living	(U5) Urban Outskirts 1	36.1	Middle
33	Midlife Junction	(L10) Traditional Living	(U8) Suburban Periphery 2	41.2	Middle
34	Family Foundations	(L10) Traditional Living	(U4) Metro Cities 2	39.0	Middle
35	International Marketplace	(L8) Global Roots	(U1) Principal Urban Centers 1	30.3	Middle
36	Old and Newcomers	(L4) Solo Acts	(U4) Metro Cities 2	37.1	Middle

37	Prairie Living	(L11) Factories and Farms	(U11) Rural 2	41.3	Middle
38	Industrious Urban Fringe	(L8) Global Roots	(U5) Urban Outskirts 1	28.6	Middle
39	Young and Restless	(L4) Solo Acts	(U4) Metro Cities 2	28.7	Middle
40	Military Proximity	(L6) Scholars and Patriots	(U8) Suburban Periphery 2	22.5	Middle
41	Crossroads	(L12) American Quilt	(U9) Small Towns	32.1	Lower Middle
42	Southern Satellites	(L11) Factories and Farms	(U11) Rural 2	37.7	Lower Middle
43	The Elders	(L5) Senior Styles	(U8) Suburban Periphery 2	73.5	Lower Middle
44	Urban Melting Pot	(L8) Global Roots	(U1) Principal Urban Centers 1	36.4	Lower Middle
45	City Strivers	(L3) Metropolis	(U2) Principal Urban Centers 2	32.3	Lower Middle
46	Rooted Rural	(L12) American Quilt	(U11) Rural 2	42.2	Lower Middle
47	Las Casas	(L8) Global Roots	(U2) Principal Urban Centers 2	25.7	Lower Middle
48	Great Expectations	(L7) High Hopes	(U5) Urban Outskirts 1	33.2	Lower Middle
49	Senior Sun Seekers	(L5) Senior Styles	(U9) Small Towns	52.5	Lower Middle
50	Heartland Communities	(L5) Senior Styles	(U9) Small Towns	42.0	Lower Middle
51	Metro City Edge	(L3) Metropolis	(U6) Urban Outskirts 2	29.4	Lower Middle
52	Inner City Tenants	(L8) Global Roots	(U4) Metro Cities 2	27.9	Lower Middle
53	Home Town	(L11) Factories and Farms	(U8) Suburban Periphery 2	34.0	Lower Middle
54	Urban Rows	(L3) Metropolis	(U2) Principal Urban Centers 2	33.3	Lower Middle
55	College Towns	(L6) Scholars and Patriots	(U6) Urban Outskirts 2	24.4	Lower Middle

56	Rural Bypasses	(L11) Factories and Farms	(U11) Rural 2	38.0	Lower Middle
57	Simple Living	(L3) Metropolis	(U6) Urban Outskirts 2	40.7	Lower Middle
58	NeWest Residents	(L8) Global Roots	(U2) Principal Urban Centers 2	25.4	Lower Middle
59	Southwestern Families	(L9) Family Portrait	(U6) Urban Outskirts 2	28.6	Lower Middle
60	City Dimensions	(L8) Global Roots	(U4) Metro Cities 2	29.2	Lower Middle
61	High Rise Renters	(L8) Global Roots	(U2) Principal Urban Centers 2	30.1	Lower Middle
62	Modest Income Homes	(L3) Metropolis	(U6) Urban Outskirts 2	35.7	Low
63	Dorms to Diplomas	(L6) Scholars and Patriots	(U4) Metro Cities 2	21.7	Low
64	City Commons	(L9) Family Portrait	(U2) Principal Urban Centers 2	24.6	Low
65	Social Security Set	(L5) Senior Styles	(U2) Principal Urban Centers 2	45.8	Low

Appendix B

Trade Area Block Group
Penetration Ratio Data

Block Group ID	Segment Code (Main Effect LIFESTYLE)	Life Mode Code (Main Effect LIFECODE)	Urban Code (Main Effect URBANCODE)	Penetration Ratio
1	41	L12	U9	0.010230179
2	41	L12	U9	0.02320442
3	28	L7	U4	0.017563117
4	36	L4	U4	0.005861665
5	22	L3	U3	0.009090909
6	14	L5	U7	0.011627907
7	22	L3	U3	0.002096436
8	22	L3	U3	0
9	1	L1	U3	0.006578947
10	1	L1	U3	0.007722008
11	3	L1	U3	0.011210762
12	36	L4	U4	0.002074689
13	22	L3	U3	0.011730205
14	9	L2	U3	0.014204545
15	22	L3	U3	0.006134969
16	22	L3	U3	0.007792208
17	55	L6	U6	0
18	63	L6	U4	0
19	22	L3	U3	0.020044543
20	55	L6	U6	0.002145923
21	36	L4	U4	0.011013216
22	55	L6	U6	0.00273224
23	63	L6	U4	0.04
24	55	L6	U6	0.001855288
25	55	L6	U6	0.007117438
26	65	L5	U2	0.018823529
27	22	L3	U3	0.00228833
28	36	L4	U4	0.003389831
29	64	L9	U2	0
30	64	L9	U2	0
31	64	L9	U2	0
32	64	L9	U2	0.002849003

33	62	L3	U6	0
34	64	L9	U2	0.00486618
35	48	L7	U5	0
36	53	L11	U8	0
37	53	L11	U8	0
38	48	L7	U5	0
39	48	L7	U5	0
40	32	L10	U5	0
41	53	L11	U8	0
42	51	L3	U6	0.013559322
43	60	L8	U4	0.004893964
44	32	L10	U5	0
45	39	L4	U4	0.002805049
46	7	L1	U7	0.010679612
47	29	L5	U8	0.011958146
48	36	L4	U4	0.006648936
49	3	L1	U3	0.010652463
50	3	L1	U3	0.02443609
51	14	L5	U7	0.012915129
52	2	L1	U7	0.051923077
53	3	L1	U3	0.016420361
54	14	L5	U7	0.023622047
55	29	L5	U8	0.014245014
56	33	L10	U8	0.016778523
57	29	L5	U8	0.012295082
58	14	L5	U7	0.014141414
59	14	L5	U7	0.021934197
60	36	L4	U4	0.007849294
61	52	L8	U4	0
62	39	L4	U4	0.004444444
63	57	L5	U6	0.000910747
64	32	L10	U5	0.008350731
65	53	L11	U8	0.002386635
66	19	L9	U3	0.005221932
67	52	L8	U4	0.001968504
68	33	L10	U8	0.005602241
69	28	L7	U4	0.002902156
70	18	L2	U8	0.001879699
71	24	L10	U5	0
72	41	L12	U9	0.002512563

73	32	L10	U5	0.0027894
74	28	L7	U4	0.001033058
75	19	L9	U3	0.004444444
76	28	L7	U4	0.003159558
77	28	L7	U4	0.012386851
78	33	L10	U8	0.007228916
79	16	L2	U3	0.027077849
80	13	L2	U7	0.021489971
81	39	L4	U4	0.013333333
82	13	L2	U7	0.013368984
83	36	L4	U4	0.007507508
84	13	L2	U7	0.012311902
85	17	L2	U10	0.367088608
86	7	L1	U7	0.285714286
87	4	L1	U5	0.344992051
88	6	L1	U7	0.4271556
89	17	L2	U10	0.275789474
90	25	L11	U10	0.284337349
91	7	L1	U7	0.118686869
92	16	L2	U3	0.034069098
93	22	L3	U3	0.077079108
94	2	L1	U7	0.218774861
95	12	L9	U7	0.067467652
96	6	L1	U7	0.053674649
97	13	L2	U7	0.052884615
98	39	L4	U4	0.013860014
99	28	L7	U4	0.010940919
100	14	L5	U7	0.0975
101	30	L5	U4	0.027850305
102	33	L10	U8	0.020733652
103	36	L4	U4	0.013867488
104	39	L4	U4	0.006479482
105	33	L10	U8	0.017910448
106	39	L4	U4	0.007181329
107	13	L2	U7	0.043367347
108	39	L4	U4	0.001683502
109	7	L1	U7	0.609090909
110	25	L11	U10	0.102222222
111	16	L2	U3	0.007751938
112	7	L1	U7	0.012578616

113	33	L10	U8	0.002721088
114	16	L2	U3	0.002523129
115	12	L9	U7	0.006479482
116	25	L11	U10	0.0234375
117	6	L1	U7	0.013422819
118	17	L2	U10	0.035781544
119	26	L12	U10	0
120	13	L2	U7	0.013320647
121	18	L2	U8	0.022508039
122	33	L10	U8	0.011331445
123	52	L8	U4	0.005366726
124	28	L7	U4	0.000892061
125	13	L2	U7	0.01369863
126	33	L10	U8	0.011764706
127	52	L8	U4	0
128	48	L7	U5	0.008350731
129	50	L5	U9	0
130	53	L11	U8	0.028985507
131	52	L8	U4	0
132	55	L6	U6	0
133	64	L9	U2	0
134	34	L10	U4	0
135	64	L9	U2	0
136	55	L6	U6	0.333333333
137	51	L3	U6	0
138	60	L8	U4	0
139	48	L7	U5	0.005037783
140	32	L10	U5	0.00591716
141	28	L7	U4	0.001033058
142	19	L9	U3	0
143	19	L9	U3	0.004444444
144	32	L10	U5	0
145	48	L7	U5	0
146	28	L7	U4	0
147	28	L7	U4	0.006036217
148	28	L7	U4	0.003745318
149	17	L2	U10	0.173210162
150	12	L9	U7	0.008178439
151	2	L1	U7	0.001814882
152	18	L2	U8	0

153	13	L2	U7	0.022222222
154	7	L1	U7	0.013953488
155	18	L2	U8	0.012232416
156	24	L10	U5	0.005586592
157	42	L11	U11	0.12327044
158	26	L12	U10	0.163339383
159	12	L9	U7	0.027543994
160	4	L1	U5	0.00995671

Appendix C

Analyses of Variance

Community Tapestry

The GLM Procedure

Number of Observations Read 115
Number of Observations Used 115

Dependent Variable: LOGPENT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	20	228.2987415	11.4149371	6.54	<.0001
Error	94	163.9827835	1.7444977		
Corrected Total	114	392.2815250			

R-Square	Coeff Var	Root MSE	LOGPENT Mean
0.581977	-25.44788	1.320794	-5.190195

MAIN EFFECT	DF	Type I SS	Mean Square	F Value	Pr > F
LIFESTYLE	20	228.2987415	11.4149371	6.54	<.0001

Multiple Range t Tests (LSD) for LOGPENT
Means with the same letter are not significantly different.

t Grouping				Mean	N	LIFESTYLE
		A		-1.8435	4	GREEN ACRES
	B	A		-2.8445	6	EXURBANITES
	B	C		-3.8247	6	IN STYLE
	B	C		-3.8259	6	PROSPEROUS EMPTY NESTERS
	B	C		-4.0334	4	UP AND COMING FAMILIES
	B	C	D	-4.2134	4	CONNOISSEURS
		C	D	-4.3587	3	RUSTBELT RETIREES
		C	D	-4.4576	4	ENTERPRISING PROFFESSIONALS
		C	D	-4.6410	7	MIDLIFE JUNCTION
	E	C	D	-4.7774	3	CROSSROADS
	E	C	D	-5.0679	8	OLD AND NEWCOMERS
	E	C	D	-5.1480	9	METROPOLITANS
E	F	C	D	-5.1787	7	YOUNG AND RESTLESS
E	F	G	D	-5.6034	9	ASPIRING YOUNG FAMILIES
E	F	G	H	-6.1511	4	MILK AND COOKIES
	F	G	H	-6.5663	5	RUSTBELT TRADITIONS
		G	H	-6.7194	6	COLLEGE
			H	-7.0261	5	HOME TOWN
			H	-7.1231	4	INNER CITY TENANTS
			H	-7.3476	5	CITY COMMONS
			H	-7.3575	6	GREAT EXPECTATIONS

Life Mode and Urban Codes

The GLM Procedure

Number of Observations Read 138
Number of Observations Used 138

Dependent Variable: LOGCODE

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	19	314.7042788	16.5633831	8.11	<.0001
Error	118	241.0963928	2.0431898		
Corrected Total	137	555.8006716			

R-Square	Coeff Var	Root MSE	LOGCODE Mean
0.566218	-28.13687	1.429402	-5.080174

MAIN EFFECTS	DF	Type I SS	Mean Square	F Value	Pr > F
LIFECODE	11	197.1228779	17.9202616	8.77	<.0001
URBANCODE	8	117.5814009	14.6976751	7.19	<.0001

Multiple Range t Tests (LSD) for LOGCODE
Means with the same letter are not significantly different.

t Grouping			Mean	N	LIFECODE
	A		-2.9905	16	HIGH SOCIETY
B	A		-3.7110	17	UPSCALE AVENUES
B	A	C	-4.0361	4	AMERICAN QUILT
B		C	-4.2155	12	SENIOR STYLES
	D	C	-5.1196	15	SOLO ACTS
E	D		-5.5241	7	FACTORIES AND FARMS
E	D		-5.6389	12	METROPOLIS
E	D		-5.6797	13	TRADITIONAL LIVING
E	D		-5.9597	13	FAMILY PORTRAIT
E	F		-6.3050	15	HIGH HOPES
E	F		-6.5065	8	SCHOLARS & PATRIOTS
	F		-7.0549	6	GLOBAL ROOTS

t Grouping		Mean	N	URBANCODE
	A	-1.8177	7	RURAL
	B	-3.3599	26	SUBURBAN
	C	-4.7774	3	TOWNS
	C	-4.9943	22	METRO
	C	-5.3749	17	PERIPHERY
D	C	-5.6022	33	CITIES
D	E	-6.5120	14	URBAN
	E	-6.7851	6	CENTERS
	E	-6.8652	10	OUTSKIRTS