CORNELL
REAL ESTATE
REVIEW

VOL. 7, MAY 2009

Retail Site Selection:

A New, Innovative Model for Retail Development

by Joshua K. Ladle and David and Duane Stiller

Retail Site Selection:
A New, Innovative Model for Retail Development

by Joshua K. Ladle and David and Duane Stiller, *Woolbright Development, Inc.*

Author

Joshua Ladle graduated from Brigham Young University – Idaho with a Bachelor of Science degree in Business Management. Following graduation, he joined Barlocker Insurance Services (BIS), a top 100 commercial insurance broker in the U.S., headquartered in Central California. As an Account Executive, he worked directly with CEOs and Senior Management to identify potential exposures to loss, develop customized solutions, and then implement them to mitigate risk. During his four years at BIS, Joshua earned the industry's Associate of Risk Management (ARM) designation, dedicated considerable time to his community by serving as Co-Chair for one of the Rotary Club's largest fundraisers and served as the youngest elected President of the Kiwanis Club. During this time he also began to realize his strong desire to move into the real estate industry. Joshua pursued this desire and applied his knowledge to form and run a multi-family housing risk management group, which provided apartment owners with industry specific insurance products and risk management techniques to protect their assets. Joshua is a 2010 candidate in the Cornell Program in Real Estate and will pursue a career in real estate finance and investment upon graduation.

Abstract

This paper provides insight into the site-selection process for retail real estate decisions. After briefly exploring current methods of retail site selection, this paper will present a new method for site selection: the "regression approach for retail site selection." This method—introduced through a case study determines a retailer's key site selection criteria based on historical data from past site selections. A developer can then extrapolate any retailer's past decisions to find attractive new sites for the retailer.

Introduction

The retail industry is rapidly changing (see Figure 1). The world's financial crisis has driven several large retailers, including Circuit City and Lines N' Things, into bankruptcy reorganization while driving other retailers, such as Goody's, Steve & Barry's, and Mervyns, completely out of business. Due to these bankruptcies and other economic conditions -- store closures caused by the tidal wave of residential foreclosures, rising unemployment, and reduced disposable income, retail REITs and other retail property owners are suffering. In this new retail environment, with a host of expansion option spanning from empty boxes to untapped raw land, retail site selection has become increasingly important to corporate real estate practitioners and the retail property developers who service them.

Figure 1: Recent Retail Headlines

"Wave of Bankruptcy Filings Expected From Retailers in Wake of Holidays" - Jan. 12, 2009, Wall Street Journal
"Discounts Fail to Save Retailers" - Dec. 31, 2008, Wall Street Journal
"S&P: More Negative for Retailers in '09" - Dec. 30, 2008, Wall Street Journal
"U.S. Retailers Face Grim Outlook" - Dec. 29, 2008, Financial Times
"Dismal Outlook for Mall Owners" - Dec. 29, 2008, Wall Street Journal
"Retailers Brace for Major Change" - Dec. 27, 2008, Wall Street Journal
"Retailers Braced for 'Horrible' Year" - Dec. 26, 2008, Financial Times
"Retail Sales Plummet" - Dec. 26, 2008, Wall Street Journal
"Retail Sales Are Weakest in 35 Years" - Dec. 5, 2008, New York Times
"Retail Faces a Long Season of Stress" - Dec. 3, 2008, Wall Street Journal
"Retail Insolvencies Expected to Rise in New Year" - Dec. 3, 2008, Wall Street Journal

Retailers are being forced to justify the existence of their network of stores. Many are coming to the realization that they grew too fast and should scale back, if not halt, new store openings for 2009 and 2010. Retailers must now determine which stores should remain open and which should be closed. For those retailers able to grow, the site selection process has become increasingly important.

In fact, today's retailers find themselves in much the same predicament as Depression Era retailers. After a cycle of retail expansion in the 1920s, retailers in 1930s found that store expansion was "slow" and that "proposed sites [were] accepted only when of obvious value and undoubted profitability" (Ratcliff, 1939). As in the past, a retail site will be chosen only if it meets all of a retailer's key criteria. No longer will factors necessary to a store's success be overlooked due to a strong economy or sheer optimism about future sales.

Retail developers will need to recognize this altered retail environment and understand the impact it will have on its retail customers. No longer will developers have the luxury of pitching a myriad of retailers on a possible site for the retailer's next store. The new list of potential anchor tenants—those considering some expansion next year—will be much shorter. As a consequence, developers' confidence in pursuing their next project will be commensurately diminished. Because retailers consider location mistakes to be among the most costly and complicated to rectify (Buckner, 1998), developers must understand that, in the near term, retailers will be more involved, more diligent, and more guarded in choosing their next site. They cannot afford a non-performing store in this difficult market. Developers will benefit greatly from an improved understanding of how retailers make their real estate site selection decisions and what criteria they use.

Corporate Retail Real Estate Decision-Making

Corporate Real Estate and an Efficient Network

Retailers are internally and externally driven to open new stores. Externally, they are faced with significant pressures from competitors to expand their network of stores. Internally, shareholders expect tangible evidence of expansion and growth (Breheny, 1988). Intrinsically, retailers seek to increase sales and market share, to better service their customer base, to hedge against the uncertainties of the market environment, and to obtain economies of scale in advertising and distribution (McLafferty & Ghosh, 1987). This also results in expansion. David Stiller, Senior Acquisition Analyst of Woolbright Development, Inc.,[1] a retail development company located in Florida, articulates the manner in which Florida retailers expand:

> Take any retailer in the state of Florida and you will find an efficient network of stores laid out to service their customer base. This network did not develop overnight. Instead stores open a few a year, here and there, in what appears to be an almost haphazard set of geographic locations. However, upon reviewing the results it appears to be an extremely efficient reaction to market conditions.
>
> A retailer services a customer base in order to make a profit. A retailer needs a certain number of customers willing to spend a certain number of dollars in their store on products before the costs of running the store are met. Exceeding those levels of sales, stores simply make greater and greater profits until a store gets so busy that another location is needed to continue to service the customers adequately.
>
> Take the following example. Everyone eats. Everyone buys food. Everyone needs to go to a grocery store to buy food. The dominant grocery chain in Florida, Publix, covers all areas of the state where you find people.
>
> The state of Florida is a growing market. Year after year, three states have consistent population growth. These include California, Texas, and Florida. The growth trend will continue. As you get more people, new stores will need to service these new comers. Remember, everyone eats and so there is always a robust network of grocery stores nearby (Stiller D. , 2007).

Author

Duane Stiller has always believed in making Florida's places better. He founded Woolbright Development, Inc. in 1983, and from the beginning he established industry knowledge, personal service and creativity as core values for the Woolbright team. Mr. Stiller built Woolbright Development into one of the fastest growing retail real estate companies in Florida, and by 2006, Woolbright was acquiring shopping centers valued at over $1 billion annually. In early 2007, Woolbright initiated a joint venture program that will power the addition of $5 billion of properties to its portfolio over the next five years.

[1] Woolbright Development, Inc. is the subject of this paper's case study.

While retailers are constantly seeking to expand their network of stores, retailers today face significant pressures and are finding expansion difficult. A rapid addition of stores during the preceding years has left many retailers questioning the efficiency of their network. Many of these retailers will need to close stores to lower operating costs. Retailers will attempt to move toward profitability and a more efficient network of stores, which is better able to service the customer base. Given the aforementioned anecdotal evidence, retailers will be opening fewer stores in 2009 and 2010 than in years past.

Large, financially able retailers will have the greatest chance of expanding in this marketplace and will devote considerable resources to their site selection process. These large retailers have considerable real estate holdings, routinely open new stores, and have sizable internal real estate departments to manage their real estate needs. The employees within these corporate real estate departments are typically trained in real estate and have significant experience dealing with brokers, land owners, and the like. These corporate real estate practitioners make site selection decisions with increasing skill and sophistication, but it is not clear that any of them are simply following a set formula. If they were following a formula, it would be easy to determine where all of the future stores would be located, but this is far from the way things work. Retail experts employed at the large retailers insist "there is no set formula," yet it has been found that much can be derived from a model based on a large number of their past decisions. In fact, such models give an excellent prediction of where future stores are needed.

The Site Selection Process and Identification of "Musts" and "Wants"

The responsibility to find and acquire suitable, i.e. profitable, retail sites is becoming more important in today's challenging economic climate. In the past, site selection "for the most part [was] a hunch-driven, hit-or-miss affair," according to Terry Meyer (1988). This, however, is not the case today. Given its increased importance, retailers need to "approach [the key dimensions of site selection] in a disciplined, systematic manner commensurate with the importance of the underlying decision" (Rabianski, DeLisle, & Carn, 2001). Companies seem to agree with this logic. Published research shows that, when selecting locations, companies follow systematic processes (Rabianski, DeLisle, & Carn, 2001).[2] By using a disciplined, systematic process, corporate real estate professionals are able to consider all available locations and identify sites which offer the greatest potential for profitability.

In years past, this systematic site selection process was seldom performed. According to Ratcliff (1939), back in 1939, only a few of the largest retail chains used "scientific" methods to identify attractive sites. Today, retailers recognize that their site location decision "is perhaps the most important decision [they] have to make" (McLafferty & Ghosh, 1987), and they should employ a "scientific" site selection methodology.

This site selection process, while different for every retailer, typically involves the same steps. An analysis of Fortune 500 firms, performed by Schmenner (1982), involved corporate plant location decisions. The study identified an eight-step sequence of incremental decisions involved in the corporate plant site selection process (see Figure 2).

Author

After graduating from Cornell University with a Master of Engineering degree in 1993, David Stiller launched his career working for several prestigious companies including JP Morgan, Credit Suisse, the Royal Bank of Canada, Rueters, and National Grid before joining Woolbright Development in 2005. Currently, he is member of the research and acquisition team focusing on locating potential opportunities. He is especially enjoying being able to apply the mathematical tools of the financial and statistical modeling worlds to commercial real estate.

[2] Rabianski, DeLisle and Carn. For additional readings, please see the articles written by Freed, S., Ettlinger, Nancy, Clay, Bradley, and Enright, M., in the Bibliography.

Figure 2: Eight Step Site Selection Process

1	The decision to seek a new site, with notification to corporate staff members involved in site selection.
2	Decisions relating to size and operational requirements for the plant under consideration.
3	Decisions relating to the design and engineering of the plant, pursued simultaneously with the location search.
4	Decisions relating to the key location criteria used in developing a "must" list (conditions which have to be met at any new location) and a "wants" list (remaining location factors that are desirable but not essential).
5	Regional location selection decision(s) to designate candidate regions using the "must" and "wants" list.
6	Decisions to include specific available, desirable sites in communities within candidate regions to form a list of alternative sites for evaluation.
7	Decisions to reduce the number of alternative sites for intensive site-specific analyses.
8	Site selection decision, initially made by division personnel with corporate approval, using results of comparative site-specific analyses.

While Schmenner's research addressed corporate plant locations rather than retail locations, many of the same steps are followed in the retail industry. Applying these steps to retail, the process starts with a decision to seek a new site, which is followed by decisions regarding the size, design, and operational requirements of the desired site. Then, the retailer must identify key location criteria through 'must' and 'want' lists. These lists are used to analyze various sites for potential fit. Finally, after the retailer performs an intensive site-specific analysis of numerous alternative locations, a site is selected and approved using comparative site-specific analyses.

Step Four of this Eight-Step Site Selection Process provides the key insight. When seeking to expand, companies produce certain criteria that are on their 'must' list and others that are on their 'want' list. Retailers operate in the same way—using their 'must' and 'want' criteria—when searching for and selecting a potential site.

Access to Retailer's Key Criteria

But which criteria do retailers place on their 'must' list and which do they place on their 'want' list? Retailers typically know their site selection 'musts' and 'wants'. Though, they may not identify them as such. Companies generally keep that information internally, with few outsiders given access to such information. "Each chain has its own formula for determining what population and demographics are needed to support an outlet, but those formulas are closely guarded secrets," explains Kristy Reynolds, a professor of marketing and management at the University of Alabama (Woodruff, 2009). Rabianski, DeLisle and Carn. For additional readings, please see the articles written by Freed, S., Ettlinger, Nancy, Clay, Bradley, and Enright, M., in the Bibliography.

A Denver-based study highlights the fact that retailers have specific criteria they use when searching for a new store site. Conducted by Charles G. Schmidt (Schmidt), Department of Geography professor at the University of Colorado at Denver, this study illustrates that site selection information primarily comes through internal sources. He derived his findings from informational interviews with the President of a Denver-based retailer. In 1962, the company started with one location and, by 1980, had forty retail locations. Through personal interviews, he identified four main site characteristics that were desired by the retailer, i.e. on its 'must' list:

1. High traffic volume
2. Maximum street frontage, wide curb cuts, and safe access to traffic in both directions
3. Parcel size, i.e. room for expansion
4. Community population threshold

In addition to these ‘musts,’ he also noted several ‘wants’:

1. Corner locations
2. Developed land rather than greenfield parcels
3. Acquisitions rather than ground-up construction

Schmidt notes that, “although substantial quantities of spatially coded data on store performance are gathered and analyzed systematically, there are no formal (written) location checklists or site evaluation documents available” (Schmidt). So, while this retailer had invested considerable time analyzing sales figures to gauge existing store performance, and obviously saw the value of empirical data in decision making, it failed to compile a list of key criteria or location checklists for real estate site selection and evaluation. Thus, while retailers certainly do have ‘must’ and ‘want’ lists for their next sites, for many retailers, these lists are found only in soft copies in the heads of those that make the site selection decisions. Hence, it would be extremely difficult for someone outside the retailer to identify the retailer’s key site selection criteria without a personal interview with its main decision maker.

Another researcher, Brubaker (2004), conducted a study to identify the site selection practices of retail tenants. The researcher conducted interviews with real estate department executives, regional directors, and research staff at many of the leading retail tenants, as well as, real estate developers, municipal planners, and brokers. Again, the ‘must’ and ‘want’ items came primarily from interviews with executives and employees of retailers. Again, this information was not generally available to the public. The ‘musts’ and ‘wants’ lists included signage, visibility, traffic counts, site size, parking, co-tenancy, proximity to other draws such as theaters, restaurants, and health clubs, demographics, i.e. population, income, education, competition and trade area. Brubaker then compiled a chart cross-referencing retailers and their ‘must’ and ‘want’ items (see Figures 3 & 4).

Figure 3: Site Selection Criteria by Retail Tenant

Tenant Type	Mainstream Electronics	Office Supply	Wholesale Clubs	Large-Format Discount Stores
Examples	Best Buy, Circuit City	Staples, Office Depot	BJ's Wholesle, Sam's, Costco	Target, Wal-Mart
Prototype Size	20,000 - 45,000 sf	20,000 - 30,000 sf	100,000-175,000 sf	90,000 to 140,000 sf Super Centers: 140,000 to 200,000 sf
Parking Requirements	6 spaces/1,000 sf (accomodates seasonal shopping)	5 spaces/1,000 sf	5.5 to 6 spaces/1,000sf	4 to 6 spaces/1,000sf
Trade Area Extent	5-10 miles	5 miles	Suburbs: 3 - 20 miles	Walmart: min: 3 miles Target: min 5 miles
Min. Trade Area Population Requirements	250,000 population	150,000 population plus 5,000 small businesses	75,000 population	100,000 to 250,000 population
Traffic/Access Requirements	Major arterial	Major arterial, high traffic	Major arterial	Major arterial (40,000+ADT)
Demographics	Growth areas, high Incomes,	High income areas, large proportion of small business activity		Higher % of high household income, especially Target. Avoids extremely high or extremely low incomes

Figure 4: Site Selection Criteria by Retail Tenant (Continued)

Tenant Type	Arts and Crafts	Bath and Linens	Bookstores	Large Off-Price Apparel
Examples	Joann's, Michaels	Linens N' Things, Bed, Bath & Beyond	Barnes and Noble, Borders	Ross, Marshalls, TJ Maxx
Prototype Size	15,000 - 30,000 sf	20,000 - 40,000 sf	20,000 - 45,000 sf	30,000 sf
Parking Requirements	5 spaces/1,000 sf	5 spaces /1,000 sf	5 spaces /1,000 sf	5 spaces /1,000 sf
Trade Area Extent Min. Trade Area Population Requirements	5-10 miles 200,000 to 300,000 population	5-7 miles 130,000 to 150,000, 10,000 to 20,000 Households with over $50,000 Income	5 miles 200,000 population, 25,000 college educated people. Significant daytime population	2-3 miles 100,000 to 150,000 population
Co-Tenancy	Female oriented stores such as bath and linens, or off-price fashion	Bookstores, off-price fashion Do not want to be by office supply, pet, restaurant, theater, gym tenants	Any high-traffic tenant that doesn't detract from image. Theaters, restaurants Do not want to be by office, pet, auto, dollar stores	High-traffic tenants. grocery stores, bath and line stores, other clothing stores.
Traffic/Access Requirements	Major arterial	Major arterial, high traffic	Major arterial, high traffic	
Demographics	High Income areas, high % of female population.	High population growth, high home ownership ratios. High Income areas	Above average incomes, highly educated areas, slightly older populations	Mid to upper income, high % of female population, high % white collar, high % of incomes $40,000 - $50,000

A Developer's Dilemma

The research done by Schmidt and Brubaker identifies valuable site selection criteria for several 'big box' retailers. What about retailers who were not a part of these studies? Brubaker's findings indicate that no two retailers use exactly the same criteria (see Figures

3 & 4). Nor could one apply the studies' data to other similar retailers and confidently conclude that one properly identified the retailers' key site selection criteria. (Though, this data is likely a good starting point.) With today's computing technology allowing the quick and efficient analysis of vast amounts of data and access to geocoded databases (McLafferty & Ghosh, 1987), much more can be done and must be done to unlock the secret site selection criteria of retailers.

The key questions boil down to: What are a retailer's key site selection criteria and how can they be identified? How can a developer use this knowledge to gain a competitive edge and survive in this new economy? To solve these problems, a developer must first understand the current site selection methods and practices used by retailers to identify favorable retail sites.

Review of Site Selection Methods

McLafferty and Ghosh (1987) deduced that no one method can be used in every situation. Buckner (1998) added that the use of numerous site evaluation methods can minimize the inherent weaknesses in any one technique. However, rather than give a provide a detailed analysis of every site selection method currently used, a cursory overview of the main approaches will be performed to provide context and to illustrate the evolution of site selection methods.

Intuition and Experience Model

For many retailers, particularly in the past, retailers made decisions based on gut instinct. Before the 1980's, the majority of large, British retailers made their location decisions based on intuition and past experience, according to Breheny (1988). However, as retailers began recognizing the critical importance of a store's location, many British retailers—and retailers throughout the world—started using more systematic and analytical forecasting techniques in the site selection process (Breheny, 1988).

Analog Model

The Analog Model constituted the first attempt at a formal retail site selection process (Applebaum, 1968). The creator of this model, William Applebaum, focused on the study of existing retail stores to identify potential retail sites. Customers of these existing stores were interviewed to determine where they lived, allowing Applebaum to define primary trade areas for these stores. He used existing store sales levels to project sales potential of future locations (Buckner, 1998). To determine the likely performance of the planned store, he performed a systematic comparison of the characteristics of the proposed store with the characteristics of the existing 'analogue' store (Breheny, 1988).

One advantage that the Analog Model has over other methods is its adaptability to assess virtually all types of retail stores. By contrast, the Gravity Model (explained next) is used primarily with supermarkets and drug stores (Buckner, 1998). However, the Analog Model has two clear weaknesses: 1) It is highly subjective and, typically, does not work well without an experienced analyst; and 2) Developing and maintaining the database through a well-trained staff has a relatively high cost (Buckner, 1998).

Gravity Model

The Gravity model is a widely used technique in retail site selection. It is derived from William J. Reilly's "Law of Retail Gravitation." In essence, it is a method of evaluating human behavior that measures the likelihood that individuals will gravitate toward a store depending on the individuals' travel distance, the travel distance to alternative stores, and the inherent drawing power of each location (Meyer, 1988). Reilly's Law can be expressed mathematically as:

$$D_{A\rightarrow B} = \frac{d}{1+\sqrt{P_B/P_A}}$$

in which:

d = distance in miles on major paved roads between two cities/towns, A and B
P_A = population of City A
P_B = population of City B
$D_{A\rightarrow B}$ = limit of City A's trading area, measured in miles along road to City B

Note: For retailers interested in analyzing the drawing power of a new store (or shopping center), they might substitute the floor space (square footage) of two stores for the city population variables in the equation above, and substitute driving time for the distance variable (Nelson, 1958).

Prime trading areas and secondary (tertiary) trading areas are then determined. Once these areas have been defined, the population is totaled for the prime trading area and a national formula created for average expenditures per family for the store type used. Next, a capture percentage is then predicted, e.g. 50%, for the prime trading area and then for each subsequent trade area, with percentages generally decreasing as the distance from the store increases (Nelson, 1958).

Nelson (1958) notes that the Gravity model provides the advantage of working well with typical, simple situations that use conservative calculations. Compared to other models, gravity models are fairly inexpensive to use: they do not require the development—and subsequent maintenance—of a store database (Buckner, 1998). Buckner (1998) points out that Gravity models use relatively few data points, e.g. population, demographic, and competitive information regarding a store's trade area. This makes it simpler than other methods. It also allows an analyst to conduct multiple "what-if" scenarios.

While deserving of praise for its strengths, Gravity Models also have considerable weaknesses. Because of its reliance on sales data, its use is limited to retail segments in which that information is readily available, e.g. supermarkets and drug stores (Buckner, 1998). Nelson (1958) also observes that this model is frequently inaccurate. The Gravity Model's inaccuracy is caused by three issues. First, the model relies on just two factors, floor space and driving time, to determine consumer shopping habits. Other factors that need to be considered include a shopping center's reputation, the number of stores at a shopping center, the quality of the merchandise, and parking availability. Second, the Gravity Model formula usually doesn't take into account the business that comes from public transportation or on foot. Third, the model assumes that a predictable percentage of sales will come from the prime trading area, the secondary trading area, and areas beyond. Given the material differences in income, ethnicity, and other factors within and among trade areas, sales percentages can rarely be predicted with accuracy.

Multiple Regression Model

The Analog Model evolved into the Multiple Regression Model. This model performs essentially the same function as the analog model. It determines the relationship between store sales and a range of characteristics, i.e. store distance, population, and competition (Breheny, 1988).

The Multiple Regression Model seeks to establish the relationship between certain independent variables and the sales variations between a group of stores (Breheny, 1988). Once an equation has been produced, it then can be used to forecast sales turnover for a proposed store by substituting values for the independent variables (Breheny, 1988). The starting equation, as defined by Breheny (1988), is:

$$\underset{\text{Dependent Variable}}{Y} = \underset{\text{Constant}}{a} + \underbrace{\underbrace{bX1 + bX2 \ldots bXn}_{\text{Store Characteristics}} + \underbrace{bXn+1 \ldots bXm}_{\text{Catchment Area Characteristics}}}_{\text{Independent Variables}}$$

where:

Y = dependent variable, e.g. store sales turnover, or money flowing to a store from an area

X = independent variables, separated into two groups: Store Characteristics (store size, parking facilities, etc.) and Catchment Characteristics (population, competition, etc.)

Multiple Regression Models hold considerable advantages over other models. Once created, regression models are simple to implement and do not require highly skilled analysts to operate and decipher (Buckner, 1998). Additionally, they allow more complex relationships to be investigated, are more flexible, and can test numerous variables relatively quickly (Breheny, 1988).

However, like the Analog Model, it also suffers from a potential methodology flaw. The selection of an appropriate analogue group, from which results are generated, is of critical importance. According to Breheny (1988), when deciding which stores to group together, the modeler must group stores according to the type of store the retailer intends to build in the future. Also, as is the case with all statistical studies, small groupings of stores must be avoided and a larger numbers of stores should be sought to achieve statistical confidence (Breheny, 1988).

Spatial-Allocation Model

Spatial-Allocation Models make use of geocoded databases and powerful computers to evaluate numerous location options and select the site that best fits corporate objectives, e.g. market share or profits (McLafferty & Ghosh, 1987). These models have only recently become popular because of technological advancements. In addition to determining the most favorable site, Spatial-Allocation Models allocate a chain's total potential sales within a market to each individual store site (Buckner, 1998).

These models "provide an efficient, powerful technique for creating decision support systems for developing location strategies" and are able to "systematically evaluate the impact of each store on the entire network of outlets in a market area" (McLafferty & Ghosh, 1987). Therefore, these models are more appropriate to use to assess the impact of a store's opening on a retailer's network of stores. By contrast, the other models fail to factor in the effect of 'one more store' within a market. Despite its utility, Buckner (1998) finds that Spatial-Allocation models are costly to develop and expensive and time intensive to perform.

Future of Site Selection Methods

As indicated above, retail site selection methods have evolved with technological innovation. Buckner (1998) notes three major trends that "will shape store location research in the next few decades." There will be:

1. Significant increase in the efficiency of collection of data and the development of databases
2. Use of advanced statistical and modeling techniques
3. Continued rapid evolution of geographic information systems (GIS)

Buckner (1998) identified a few statistical techniques that hold the great promise for increased use in locational research, including Neural Networks and Chi Square Automatic Interaction Detection (CHAID). As was predicted by Buckner in 1998, retail site selection methods have continued to advance and evolve over the last decade. Consistent with Buckner's three major trends and the advanced statistical techniques, Stiller created a new retail site selection method. This model constitutes a revolutionary new way to determine a retailer's key site selection criteria based on historical data of past site selections.

A Look at History and Predictability

Retailers, like consumers, tend to make similar decisions again and again. Given the same circumstances, they will act in much the same way as they acted before. It is when we begin to analyze extensive data sets of past decisions that we begin to see trends in the variables underlying the decisions.

Through the Woolbright Development case study, it will be shown that by analyzing the sites that retailers have selected over the past 15 years, developers can, with a high level of accuracy, identify the key site selection criteria of a given retailer. They can then use that data to evaluate available parcels of land—or existing shopping centers—and select the most attractive site for a retailer based on the retailer's historical site selection criteria. This capitalizes on the assumption that future selections will mirror the decision attributes of previously selected stores.

Case Study: Woolbright Development, Inc.

History and Background

Located in Boca Raton, Florida, Woolbright Development, Inc. is one of the top 10 private owners of retail real estate in Florida. Its portfolio includes more than 30 projects totaling nearly 5 million square feet of retail and mixed-use real estate (see Figure 5). It is ranked as the 7th largest open air retail center owner in Florida, as measured by Gross Leasable Area (GLA) (see Figure 6).

Founded in 1983, Woolbright concentrates on the six major metropolitan areas of the state: Orlando, West Palm Beach, Ft. Lauderdale, Miami, Tampa, and Jacksonville. It is involved in nearly every aspect of retail real estate, including Acquisitions, Development, Property Management, Leasing, and Market Research.

Since its inception, Woolbright has collected data on every shopping center with over 45,000 square feet in the six major metropolitan areas of Florida, amassing a database of over 2,300 shopping centers. Key data, such as vacancy levels, ownership, tenants, sales histories, and construction dates, have been pulled from various sources, including annual shopping center visits. Some have called this database, “one of the most powerful retail databases of its kind.” It is this database that has allowed Woolbright to explore and analyze Florida retailers’ past site selection decisions.

Figure 5: Woolbright Development, Inc. Shopping Centers

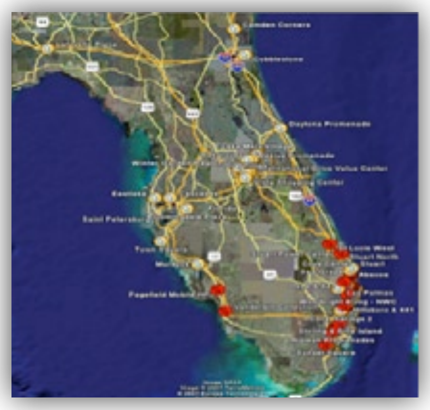


Figure 6: Woolbright Development, Inc. Ranking

Company Name	National Retail Portfolio			Florida Retail Portfolio		
	GLA (Million SF)	National Ranking	Centers	GLA (Million SF)	Florida Ranking	Centers
Developers Diversified Realty	105.6	2	458	13.2	1	83
Kimco Realty Corp	129.2	1	733	12.9	2	88
Equity One	21.6		192	10.0	3	87
Sembler	14.3		69	9.0	4	49
Centro Watt / Newplan Excel	95.7	3	596	6.7	5	40
Regency Centers	53.0		401	6.2	6	52
Woolbright Development	**6.0**		**40**	**6.0**	**7**	**40**
Weingarten Realty Investors	56.5		321	5.8	8	31
Ramco-Gershenson	14.8		83	4.0	9	24
Inland Real Estate Group*	62.1		472	1.9		15

Retail centers > 45,000 SF excluding single tenant properties and regional malls
** exclude properties owned by Inland Real Estate Exchange Corporation*

Solution Sought

Woolbright Development understands the fact that, “real estate developers make few decisions that are as crucial as selecting the site of their next development” (Brubaker, 2004). In the fall of 2007, David Stiller (2007), Senior Acquisition Analyst at Woolbright, approached the Master of Engineering Program in Cornell University’s School of Operations Research and Information Engineering with a project. The project had two main goals:

1. Study the efficient network of retailers in Florida based on historical openings to determine the criteria under which a new store enters the network
2. Develop a model to predict future expansions by the retailers

For the research team[3], Stiller (2007) laid out the dynamics of the retail industry and how Florida retailers had developed an efficient network of stores to service their customer base. He provided to the research team the additional following information:

[3] The Cornell research team included Haroun Al-Mishwit, Joanna Antisell, Sangmi Je, Kamil Tazi, and Jun Wang.

Current and Historical Retail Locations

Woolbright Development has a database [of] every major shopping center in several major markets in the state of Florida. The database includes all of the major retailers, their location, their size, the year they opened, and more.

Future Retail Locations

Woolbright has data on likely future store openings. Several sites can be tested to determine if these future sites meet historical criteria for past store openings.

Current and Historical Demographics

Woolbright has subscriptions to demographics data. Additional sources of demographics data also exist. Given a set of geographic coordinates, demographic reports can be generated for a given area. These include information about the population. Traditionally, retailers review demographics in one, three, and five mile radii from a given location.

Secondary demographics data is also available. Population can be broken down by where people live or work, various age categories, or family size. Income can be measured by median income, average income, or number of people in various income brackets. This additional level of detail might provide greater insight into specific targeted customers.

Projected Demographics

Demographics subscriptions also include five year futures estimates. These estimates can be used to predict future growth and future store locations.

Conclusion

Florida is a growing area. As people move into an area, retail follows. As more people move into an area, the area can support more retail space…

…As the population grows, the network of stores grows. What is needed is to find the next area these stores would be willing to locate. We know what the conditions were like when the stores opened. We should be able to determine the criteria that goes into the decision to locate each chain.

With a better understanding of what Woolbright Development wanted, the research team summarized its thoughts about the project, "Fundamentally, the challenge in this situation is one of information asymmetry. Retailers have internal methods for selecting sites. Woolbright Development is only profitable if the firm consistently selects properties which will be in high demand in the future. Therefore, Woolbright needs this information, that the retailers closely guard, to remain profitable, especially as more and more competitors enter the retail development space in a high population growth area like Florida. This information gap, although a major hurdle, is not impossible to close. Reverse engineering retailers' decision-making processes, if conducted as accurately as possible, is the best method to close this gap with the available information" (Al-Mishwit, Antisell, Je, Tazi, & Wang, 2008).

Methodology[4]

Existing Models and Past Research

The research team began the project by examining current models and methods for retail site selection. The team found that early models focused on the size of the store or store sales figures as dependent variables. The two most common models were the gravity model and the regression model. Because of the flaws and difficulties inherent in these models (as previously identified), the research team chose to explore other models.

Data Collection

With access to Woolbright Development's database of Florida shopping centers, as well as extensive demographic data, the research team selected a set of retailers for this study. This set, consisting of eight retailers, was divided into four main categories. In creating the categories, the team assumed that retailers in similar sectors would have similar site selection criteria (see Figure 7).

Figure 7: Retail Categories and Retailers Studied

Discounter	Home Improvement	Grocery	Office Supply
• Walmart • Target	• Home Depot • Lowes	• Publix	• Office Depot • Office Max • Staples

After identifying the retailers for the study, the team then obtained the location (latitude and longitude) and opening year of each of the retailers' Florida stores, i.e. the retailers' network of stores. Data consisting of influential site selection factors, including mean and median income, population, mean and median age, and race, were then obtained for each of the locations. In order to model trade areas[5], this data was retrieved for each of the stores at varying radial distances of 3, 5, 7, and 10 miles. The team noted that "due to the infrequent sampling of U.S. census reports, [this data was] only obtained for years 1995, 2000, and 2007."

Data Mining

Initial adjustments to the pulled data were required. For instance, a retailer's store that opened in 2002 would not have population data for that year since data was only pulled in 1995, 2000, and 2007. Therefore, the team used straight line interpolation between the years during which data was available to solve this problem. Mean and median incomes were also interpolated in the same manner, and were also adjusted with a constant inflation rate of 3%.

---

[4] Due to the space limitations of this publication and for the reader's convenience, the author has taken the liberty of summarizing the Cornell University research team's 97-page report, "Woolbright Development: A Retail Site Selection Study". The data and findings in this summary are those of the research team. All footnotes, unless otherwise noted, are to the research team's work.

[5] Trade area is the geographic area that provides the vast majority of the steady customers necessary to support a shopping center. Trade area boundaries are determined by a variety of factors, including shopping center type, accessibility, physical barriers, location of competing facilities, and driving time and distance (Kramer et al, 2008).

Anomalies

The Woolbright Development database consists of numerous sites built before 1990. Because it was assumed that retailer strategies and their site selection processes change slightly over time, and significant data was missing for these older properties, for this study, the research team considered only relatively new data (within the last fifteen years).

Initial Assumption

The radial distance away from each site varies based on retail sector. For example, all modeling for Publix Super Markets (the dominant grocery store in Florida) was performed using a radius of 3 miles since consumers are generally less willing to travel past that distance for basic weekly necessities such as food. In contrast, all modeling done for Home Depot and Lowes was done using a 10-mile radius since consumers will travel further to purchase infrequent products such as those found at home improvement retailers.[6]

Raw and Synthetic Explanatory Variables

The four raw input factors that were used to predict retailer site selection behavior were:

1. Mean and Median Income: These are Important variables because they are directly correlated to the amount of disposable income that the surrounding population possesses. Thus, aggregate retail sales volume can be predicted.

2. Population: This is likely the most important indicator of future retail potential.
3. Mean and Median Age: These factors have a significant influence over sectors where the primary customer base is comprised of homebuilders, such as Lowes and Home Depot. In addition, the age of those that purchase durable goods (electronics, computers, furniture, etc.) will also likely be deemed a significant factor to retailers that fall within the "Discounters" and "Office Suppliers" categories.

4. Race: This factor is often assumed to be associated with and strongly correlated with income.

The two synthetic input factors that were incorporated into the modeling analysis were:

1. Population Growth: This factor is considered by retail expansion strategists when calculating the future retail potential of an area. The population percentage increase from 2007 to the projected population in 2012 was used as a forward looking element for each existing store site.

2. Competitor Intensity: This factor seeks to capture the influence of an existing competitor site near a given site.

[6] McLafferty and Ghosh, in *Location Strategies for Retail and Service Firms*, estimated the trade area of a supermarket –from which it draws most of its customers—to fall within a 1.5 mile radius from the store. They also concluded that department stores and shopping centers have larger trade areas, i.e. greater radii.

Response Variable

The research team chose to use the "intensity" (a weighted average of retail outlets in a given area based on their distance from the site in question) as an indicator of the attractiveness of an area. Because it is essential to understand how a retailer's intensity is derived, an extract from the final report describing the "intensity" parameter is described below.

> Creating an intensity parameter for a site is similar to creating a score for a site based on its demographic characteristics. Creating a relationship between the demographics of a candidate site and an expected intensity given those demographics is simple and straightforward. Simply calculate the current intensity at a candidate site and compare it with the expected intensity given that site's demographic and synthetic profile.
>
> Constructing such a parameter can be accomplished in a variety of ways. All of the distance calculations, site intensity, and competitor intensity calculations were executed using a computer script written in Perl. The following is a brief explanation of the steps that were taken to create the intensity parameter which served as the response variable for this study.

- Calculate site to site distance for every location in each of the 8 retailers' Florida networks contained in the pilot study
- Define the appropriate trade area for the retail category for which the intensity is being calculated
- Calculate intensities for every location in each network, starting from a value of 1, based on the following weighting scheme:
  - Grocery Retailers (Publix)
    - o For a Publix store within a 0.5 mile radial distance from and opened prior to the Publix site undergoing the intensity calculation, add 1 to the value for site intensity.
    - o For a Publix store between 0.5 miles and the edge of the predefined trade area for (and opened prior to) the site undergoing the intensity calculation, add a value of: 1/(0.5+d)~t o the total value for site intensity, where d = distance from store to evaluated site.
  - All Other Retailers (Discounters, Office Supply, Home Improvement)
    - o For a same-retailer store within a 1.0 mile radial distance from and opened prior to the site undergoing the intensity calculation, add 1 to the value for site intensity.
    - o For a same-retailer store between 1.0 mile and the edge of the predefined trade area for (and opened prior to) the site undergoing the intensity calculation, add a value of: l/d to the total value for site intensity, where d = distance from store to evaluated site.

Figure 1 below depicts the above described weighting scheme graphically using Walmart as an example.

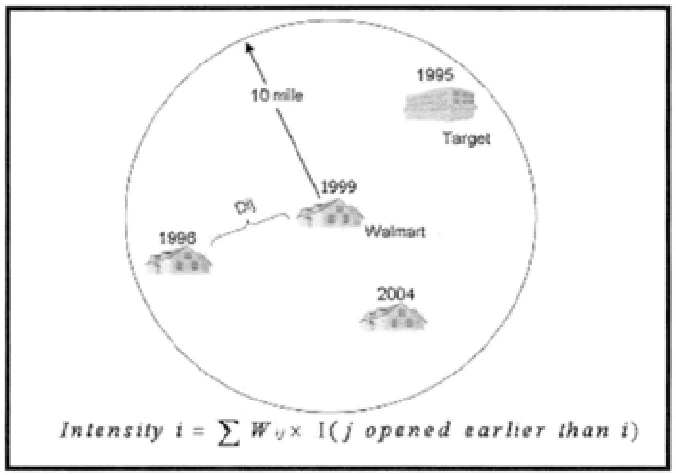


Figure 1: Weighting scheme for Walmart – 10 miles radius

The idea depicted in Figure 1 builds the foundation for all the modeling infrastructures in this study. Without it, utilizing any input-output modeling technique would be impossible. Thus, it is important to understand all of the subtleties that are involved in the calculation of site intensity. In the Figure above, the Walmart undergoing the intensity analysis was opened in 1999. Therefore, in the intensity calculation for this site, only the Walmart opened in 1996 at a distance of Dij away from this site is counted simply because it was already there when Walmart expansion strategists decided to build a new location at the current site; hence, this existing site was included in their decision making process. In contrast, the Walmart opened in 2004 was not part of the decision making process when Walmart decided to expand its network at the current site in 1999 and, therefore, is not counted in the intensity calculation.

Furthermore, a Walmart competitor, Target, is present in this specific site's trade area. The team deliberated including competitors within the intensity calculation, primarily because retailers will often fail to differentiate between sales from a given location being cannibalized by another location of their own or declining due to the presence of a competitor. Despite this mentality, the team chose to create a synthetic input factor ('Competitor Intensity,' which was discussed previously) to capture the effect of competitors. In the example above, the opening year of the competitor was prior to that of the site in question. Thus, it is counted as a competitor based on the same weighting scheme which was defined above for same-store site intensity and used as an explanatory variable in all of the modeling infrastructures utilized throughout this study (Al-Mishwit, Antisell, Je, Tazi, & Wang, 2008).

Pass/Fail Model

This model is used for locations where there are currently no other stores of a particular retailer in that trade area, e.g. there are no other Home Depot stores within a 10-mile radius. For such a location, a determination would need to be made on whether the area can support one store. If the site falls within the range of the minimum and maximum threshold values of population, mean/median income, mean/median age, competition, and white/black percent of existing stores, then the site "passes" and a store should be considered at that location. If some of the demographic numbers fall outside the ranges, then the site "fails" and that particular retailer should not be considered for that location.

Pass/Fail Implementation Using Minitab Software: The research team analyzed all of the existing stores sites for each retailer using Minitab. The software calculated the mean, standard error and standard deviation, as well as the minimum, maximum, median, 1st quartile, and 3rd quartile for each demographic data point, i.e. population, mean/median income, etc. This was done at all the radii specific to each store and then interpolated back to the year that the store was opened. See Figure 8 for the results for Target at a trade radius of 7 miles. A retailer can then use this table to construct the pass/fail model.

Figure 8: Descriptive Statistics for Target at 7 Mile Radius

Variable	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Population	285268	19903	179127	22774	152760	249581	370876	822201
Median Income	44485	829	7462	27029	40363	43864	47657	70263
Mean Income	59717	1188	10691	40463	52223	58098	64220	99609
2007 Median Age	39.744	0.590	5.310	32.912	35.698	38.033	43.135	61.686
2007 Mean Age	39.493	0.476	4.281	34.108	36.103	38.208	42.488	55.541
White Percent	0.7887	0.0131	0.1180	0.4564	0.7037	0.8023	0.8760	1.0000
Black Percent	0.1272	0.0103	0.0927	0.0021	0.0623	0.1043	0.1733	0.4556
Projected Population	0.1036	0.0066	0.0593	0.0066	0.0603	0.0911	0.1396	0.3266
Competitor	0.8016	0.0723	0.6508	0.0000	0.2341	0.7366	1.2519	3.0828

Interpretations: This pass/fail model for Target at a trade radius of 7 miles requires all the potential site's demographic variables to fall equal to or above the minimum column and to fall below or equal to the maximum column. A site that falls within the range for all demographic variables is a "pass." In some instances, such as the minimum population of 22,774 in a 7-mile radius, which seems quite low, a more reasonable test might be to use the 1st quartile result of 152,760 as a tighter check.

Regression Models[7]

Model Background and Mechanics: The Backward Elimination Method was used in this study to test for significant correlations between input and output observations. This method begins with a 'maximum model,' or in other words, a model that contains every possible explanatory variable. Variables are then removed based on a statistical hypothesis test at a predefined level (p-value[8] of 0.05 was used in this study) until no additional variables can be removed while still adhering to the limitations of the hypothesis test.

Results and Interpretations: The team's regression results varied in quality between different sectors and different retailers with R^2 values[9] that ranged from a disappointing 0.20 to a very impressive 0.80. On average, the R^2 value for all regressions performed was 0.42, which is reasonable considering many of the likely-considered factors in the site selection decision-making process were omitted from this model due to lack of data. Also, because much of the site selection decision is subjective and driven by other factors (such as human experience), it should not be expected that any model would perfectly explain the variability of site attractiveness.

[7] Regression can be defined as any process which attempts to minimize the difference between predicted values and observed values using a least squares function, using a combination of predefined model parameters.

[8] The P-value is the smallest significance level at which the null hypothesis would be rejected for a 2 alternatives test.

[9] In a regression, R^2 is the fraction of the sample variation in y, the response variable, that is explained by x, the explanatory variable.

Despite the issues discussed above, the regression results did validate the importance of certain factors in the site selection decision-making process and bring to light other factors, which were historically not thought to be important (see Figure 9)

Figure 9: Linear Regression Results

Category	Retailer (trade area)	'Best Fit' Model	Best Fit R^2	'Most Consistent' Model	Most Consistent R^2
Grocery	Publix (3 Mile)	Population White Percent	0.48	Population White Percent Median Income	0.43
Home Improvement	Home Depot (10 Mile)	Population Competitor	0.80	Population	0.74
	Lowes (10 Mile)	Population	0.23	Population	0.23
Discounters	Walmart (10 Mile)	Population Competitor Mean Income White Percent	0.55	Population Competitor	0.46
	Target (10 Mile)	Population Competitor	0.45	Population Competitor	0.45
	Walmart (7 Mile)	Population Competitor Projected Population	0.46	Population Competitor	0.35
	Target (7 Mile)	Population Competitor	0.36	Competitor	0.26
Office Supply	Office Depot (7 Mile)	Population Competitor	0.53	Population	0.47
	Office Max (7 Mile)	Population Mean Income Projected Population	0.56	Population Mean Income Projected Population	0.56
	Staples (7 Mile)	Population White Percent	0.29	Population	0.20

Neural Networks[10]

The team then looked to the Neural Networks (NN) model to analyze the data sets. "The idea of NN modeling is common among researchers, scientists, and professionals who typically work with large data sets containing complex patterns of embedded information. Within the retail real estate development practice, the idea of using a NN is slightly foreign primarily because datasets (store locations or associated demographic information) are typically no larger than the number of stores in a specific retailer's network and cannot provide enough information for a successful NN training process. However, as more information becomes available to developers via improvements in technology and record keeping, NN's will become more useful in dissecting the massive amounts of data which, if synthesized correctly, will give developers a competitive edge" (Al-Mishwit, Antisell, Je, Tazi, & Wang, 2008). Fortunately, the data set of the eight retailers contained a sufficient number of stores, so the research team was able to test the NN structure.

NN modeling was chosen by the team because it was felt that a simple linear regression model was not sophisticated enough to capture the complex patterns of

[10] In practical statistics terminology, a NN is a non-linear statistical data modeling tool which can be used to model complex relationships between inputs and outputs and to find patterns in data. The NN infrastructure consists of an interconnected group of artificial 'neurons' and processes information using a connectionist approach to computation. This connectionist approach is what one is referring to when referencing a NN's ability to find complex patterns in data. A key characteristic that is indicative of a NN is that of the adaptive ability of the system to change its structure based on the relationships in a given data set.

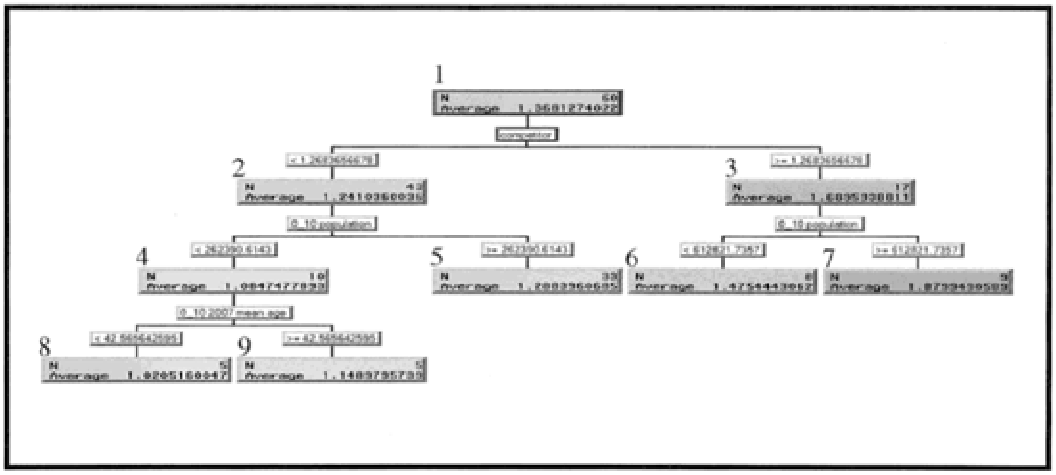
the data set. Because an in-depth review of the team's NN methodology, as well as the accompanying science behind NN, is beyond the purview of this article, readers interested in reading more about the subject should access the complete report. Suffice it to say, the research team used the NN to identify high correlation effects between some or all of the input variables and the response variable.

Decision Trees[11]

The third model the team used was the decision tree method, also called a regression tree. The leaves of the "regression tree" contain the mean values of the response variables, which can be interpreted as predicted values of the response variable. The CHAID (Chi-Squared Automatic Interaction Detection) method was used to construct the decision tree.

The decision tree is interpreted upside down as it begins with the 'root' node and then branches out, finally stopping at the 'leaves' of the tree. At each 'splitting' point, an explanatory variable, which bests splits the response variable, is selected from the pool of all possible input variables. Then, a search is conducted for the optimal value of the chosen explanatory variable on which to split the response variable into two homogenous groups. The tree 'ends' when there are no more explanatory variables that are found to significantly decrease the splitting criterion statistic any further (see Figure 10).

Figure 10: Target 10 Mile Decision Tree Example



The decision tree above is for Target at a 10-mile radius and was constructed using data from sixty Target store sites. The most significant criterion upon which Target store locations seem to be based is the number of competitors at the site, as evidenced by the first split of the tree. The next most significant criteria would be population followed by mean age.

Model Comparison

The three models employed by the research team, along with their crucial differences, are summarized in Figure 11.

[11] The goal of the decision tree methodology is to produce a model that will generalize the patterns in the existing data and that will perform well on new data. Decision tree modeling is based on performing a series of if-then decision rules that form a series of partitions to sequentially divide the response values into a small number of homogenous groups, forming a tree-like structure. Each split is performed on the values of one of the explanatory variables that best partitions the response values.

Figure 11: Primary Model Differences

	Complexity of Model	Quantity of the Data Required	Explanatory Power
Decision Trees	Low	Moderate	Low
Neural Networks	High	High	Moderate
Regression Models	Moderate	Low	High

A Perfect Model?

The research team concluded that, based on the advantages, disadvantages, and implementation issues of each model used, there is no one perfect, all inclusive method that can be used to make retail site selection decisions. Rather, the models used in this study should be used on a sector, retailer, and possibly even region-specific basis.

Validation and Empirical Results

After analyzing the eight retailers' networks of Florida stores, the team turned to the current candidate sites provided by Woolbright Development. With the data that was gathered in the study, the team was able to identify sites that were statistically an attractive fit.

These results are found in Figures 12-16. The figures contain several of the candidate sites under consideration by Woolbright Development and the retailers who should be approached for lease negotiations. The results are based on the site's demographic profile and the current intensity of each retailer at the site.

Figure 12: Attractive Publix Candidate Sites

Area Name	Extra Supporting Power by Regression	Supporting Power by Neural Network	Supporting Power by Decision Tree
0204-101	1.05	1.28	1.01
0204-102	1.06	1.30	1.01
0601-043	1.10	1.10	1.15
0602-082	1.11	1.13	1.10

Figure 13: Attractive Lowes Candidate Sites

Area Name	Supporting Power by Regression	Supporting Power by Neural Network	Supporting Power by Decision Tree
0004-078	1.09	1.04	1.05
0004-081	1.08	1.05	1.07
0004-062	1.08	1.07	1.10
0201-065	1.05	1.06	1.00
0004-080	1.05	1.01	1.03
0201-064	1.05	1.04	1.00

Figure 14: Attractive Target (7 Mile) Candidate Sites

Area Name	Supporting Power by Regression	Supporting Power by Neural Network	Supporting Power by Decision Tree
0007-098	1.23	1.28	1.35
0301-050	1.14	1.06	1.14
0602-082	1.03	1.07	1.05
0703-014	1.10	1.14	1.20

Figure 15: Attractive Walmart (7 Mile) Candidate Sites

Area Name	Supporting Power by Regression	Supporting Power by Neural Network	Supporting Power by Decision Tree
0004-051	1.05	1.33	2.49
0004-062	1.17	1.38	2.59
0004-078	1.49	1.60	2.61
0004-080	1.18	1.35	2.40
0004-081	1.25	1.43	2.49
0204-113	1.13	1.26	1.38

Figure 16: Attractive Office Depot Candidate Sites

Area Name	Supporting Power by Regression	Supporting Power by Neural Network	Supporting Power by Decision Tree
0004-062	1.22	1.35	1.15
0004-069	1.36	1.45	1.18
0004-081	1.01	1.06	1.09
0602-082	1.07	1.18	1.08
0605-061	1.29	1.41	1.09
0605-089	1.31	1.46	1.06

Note: The candidate sites were identified as attractive sites by all three modeling infrastructures for each respective retailer. As the team explained, "Sites were identified as attractive if the difference between a site's predicted intensity and current intensity for a given retailer was greater than one, meaning that based on the sites demographic profile and its proximity to other similarly branded sites, another store can be supported there. This difference between predicted intensity and current intensity was termed a site's 'Extra Supporting Power'" (Al-Mishwit, Antisell, Je, Tazi, & Wang, 2008). Additionally, many sites (not included in this paper) were only identified by two models or one model as being attractive sites. This does not mean that a retailer should not be approached regarding a certain potential site. A candidate site that is identified by all three models is the best fit for a retailer, followed by a site identified by two models, followed by a site identified by only one model.

Future Considerations

Upon completion of this project, the research team identified several site selection factors which influenced the decision-making process, but which were not considered due to data limitations. These factors could be used in future models to better determine attractive sites for retailers.

1. Location and Accessibility Features
   - Visibility: Key feature for the majority of retailers.
   - Accessibility: Essential that shoppers' movements are expedited to and from the center.
2. Demographic Features
   - Education: Historically, highly educated populations have different shopping traits than populations with lower levels of education.
3. Site Configuration Features
   - Constructed features: Data related to location sizes, shapes, frontage width and depth.
   - Natural features: Data referencing topography, soil, drainage, and vegetation.

4. Financial Characteristics
    - Potential profitability of an area depends greatly on the costs associated with operating there, e.g. high store rental rates, high local taxes, etc.

Implication for Retail Developers Now…and in the Future

In their final report to Woolbright Development, the research team summarized their study and results as follows:

> As big box retailers use increasingly more data-driven methods to choose locations for their stores, real estate development firms can gain an advantage over competitors by using more sophisticated data analysis techniques. By examining historical data, the team has developed models to help predict whether a retailer would be willing to put a new store at a candidate site. This information will allow Woolbright to evaluate sites and choose which new locations to purchase or redevelop, and specifically which retailers to target for each site (Al-Mishwit, Antisell, Je, Tazi, & Wang, 2008).

While Woolbright Development certainly has a sizable lead over other retail developers in the Florida market, they do not have a monopoly on the data necessary to create such valuable, predictive models. This data is available to any developer who recognizes the value of it and seeks to acquire it. By obtaining and using the needed data, a retail developer would be able to significantly mitigate its risk in either purchasing a vacant piece of land or in acquiring an existing shopping center because it would already know which retailers would want to be at that site.

In the short-term, it would appear that few opportunities exist for retail developers given the declining retail market. However, not all retailers are downsizing and closing stores. The Wall Street Journal reported that Best Buy Co., Costco Wholesale Corp., Au Bon Pain Inc., and Family Dollar Stores are all seeking to expand. It reported that Family Dollar, a chain of 6,600 stores, plans to open 200 stores during its fiscal year ending in August (Hudson, 2009). Additionally, it reported that Aldi, a German store chain, is expected to open 75 U.S. stores this year (Rohwedder & Kesmodel, 2009). The strong retailers are looking to expand their network of stores, and will naturally be going through the site selection decision-making process internally. Retail developers, who take the time necessary to replicate the Stiller retail site selection method will find more than enough development opportunities servicing the needs of these strong, expanding retailers.

New store openings will not be the only opportunity for retail developers. Many retailers are also closing stores throughout the nation. Retailer Goody's, which is currently going through the liquidation process, will be moving out of 282 stores (Talley, 2009). Another store, Office Depot, is reported to have plans to close 112 of its 1,275 North American stores. Other retailers like Steve & Barry's, LLC and Mervyn's, LLC, are also closing stores as they liquidate in bankruptcy court (Hudson, 2009). And KB Toys, Ann Taylor, Sears, Talbots, Circuit City (Woodruff, 2009), and Macy's (Wohl & Chaudhuri, 2009) have all announced the closures of some stores. This list merely scratches the surface of future store closures. The International Council of Shopping Centers projected back in October 2008 that 148,000 stores would close in 2008 and 73,000 stores would close in the first half of 2009 (Woodruff, 2009).

These store closures, especially of anchor stores, will cause smaller retailers to leave shopping centers because decreased foot traffic. With the 1,136 store closures that have already been announced or are expected to closed because of bankruptcy and liquidation (Bodamer, 2009), it's likely that shopping malls will see fewer shoppers this year. This, in turn, will lead to decreased sales for the retailers left behind and decreased cash flows for the property owners. Financially weak owners with significant debt service may have no choice but to dispose of these troubled assets. A developer, using the Stiller Retail Site Selection Method, can capitalize on the acquisition of these troubled assets. By knowing which retailers are looking to expand, and identifying their key internal site selection criteria, a developer can quickly target the appropriate retailers with an attractive site.

Conclusion

The retail industry is changing, and those that want to succeed in this new environment must recognize this and adapt. With retailers investing more time and money into site selection, developers must use every available resource to identify appropriate retail sites. More research will be done and double-checked to find each new site to ensure that the selected site meets all the 'musts' on the retailer's list, and hopefully some of the 'wants.' Retailers in this new economy cannot afford to pick the wrong site. "In retail, the cost of one bad store is huge" (Misonzhnik, 2006). Not only is the cost enormous, but the risk as well. "As each investment becomes greater so does the risk attached to each location decision" Breheny (1988). In today's real estate climate, retailers are experiencing this increased site selection risk.

Developers need not wait on the sidelines until the retail market recovers. There are too many current opportunities available if you know how to find them. Developers that continue to operate with the antiquated site selection methodologies will be left behind, while developers that embrace technology and find alternative ways to service retailers will not (Stiller D. , 2009). The Stiller retail site selection method, while not the 'perfect' solution, will provide those developers who utilize its predictability functions the chance to move past their competition and reap profits, even in this tough retail climate.

The research team summed it up when they declared, "The future of the retail real estate industry will be written by those firms who perfect the sound and efficient management and interpretation of large amounts of data" (Al-Mishwit, Antisell, Je, Tazi, & Wang, 2008). It has been shown in this paper that the data and process are available. The successful retail real estate developer will be the one who recognizes the need for the data, gathers the data, and deciphers the data to extract one of the most guarded retailer secrets -- their key site selection criteria.

Bibliography

Al-Mishwit, H., Antisell, J., Je, S., Tazi, K., & Wang, J. (2008). *Woolbright Development: A Retail Site Selection Analysis*. Ithaca.

Applebaum, W. (1968). The Analog Method for Estimating Potential Store Sales. In C. Kornblau, *Guide to Store Location Research*. Reading, Mass.: Addison-Wesley.

Bodamer, D. (2009, January 16). *Circuit City to Liquidate*. Retrieved January 24, 2009, from Retail Traffic: http://blog.retailtrafficmag.com/retail_traffic_court/2009/01/16/circuit-city-to-liquidate/

Breheny, M. J. (1988). Practical Methods of Retail Location Analysis: A Review. In N. Wrigley, *Store Choice, Store Location and Market Analysis* (pp. 39-86). London: Routledge.

Brubaker, B. T. (2004). *Site Selection Criteria in Community Shopping Centers: Implications for Real Estate Developers*. Cambridge.

Buckner, R. W. (1998). *Site Selection: New Advancements in Methods and Technology*. New York: Chain Store Publishing Corp.

Enright, M. (1995). *Regional Clusters and Economic Development: A Research Agenda*. Cambridge: Harvard Business School.

Ettlinger, N., & Clay, B. (1991). Spatial Divisions of Corporate Services Occupation in the United States. *Growth and Change* , 36-53.

Freed, S. (1989, Jan/Feb.). Factories of the Future: Choose New Locations Carefully. *Industrial Development*.

Hudson, K. (2009, January 7). Struggling Retailers Press Struggling Landlords on Rent. *Wall Street Journal*.

Kramer et al, A. (2008). *Retail Development, Fourth Edition*. Washington D.C.: ULI-the Urban Land Institute.

McLafferty, S. L., & Ghosh, A. (1987). *Location Strategies for Retail and Service Firms*. D.C. Heath and Company.

Meyer, T. G. (1988). Site Selection VS. Site Evaluation: Techniques For Locating Retail Outlets. *Real Estate Issues* , 13 (Spring/Summer), 25-28.

Misonzhnik, E. (2006, August 1). *Site Seeing*. Retrieved December 2008, from Retail Traffic: http://retailtrafficmag.com/mag/retail_site_seeing/

Nelson, R. L. (1958). *The Selection of Retail Locations.* New York: F.W. Dodge Corporation.

Rabianski, J. S., DeLisle, J. R., & Carn, N. G. (2001). Corporate Real Estate Site Selection: A Community-Specific Information Framework. *Journal of Real Estate Research* , 165-198.

Ratcliff, R. U. (1939). The Problem of Retail Site Selection. *Michigan Business Studies* , IX (1), 1-88.

Rohwedder, C., & Kesmodel, D. (2009, January 13). Aldi Looks to U.S. for Growth. *Wall Street Journal*.

Schmenner, R. (1982). *Making Business Location Decisions*. Englewood Cliffs, NJ: Prentice Hall.

Schmidt, C. G. (n.d.). Location Decision-Making Within a Retail Corporation. 60-71.

Stiller, D. (2009, January 6). President. (J. Ladle, Interviewer)

Stiller, D. (2007, November). Proposal for Study for Woolbright Development.

Talley, K. (2009, January 7). Retailer Goody's Plans to Liquidate. *Wall Street Journal*.

Wohl, J., & Chaudhuri, R. R. (2009, January 8). *Macy's to Close 11 Stores, Cuts Forecast*. Retrieved January 9, 2009, from Reuters UK: http://uk.reuters.com/article/companyNewsMolt/idUKTRE5071CA20090108

Woodruff, C. (2009, January 4). *'Big Box' Retailers*. Retrieved January 9, 2009, from Montgomery Advertiser: http://www.montgomeryadvertiser.com/apps/pbcs.dll/article?AID=2009901040302