Retail Site Selection And Geographic Information Systems

Article *in* Journal of Applied Business Research · September 2011

DOI: 10.19030/jabr.v11i2.5874

Jeff Tayman
University of California, San Diego

Retail Site Selection and Geographic Information Systems

Dr. Jeff Tayman, Senior Demographer, San Diego Association of Governments
Dr. Louis Pol, Kiewit Professor of Marketing, University of Nebraska at Omaha

Abstract

Geographic Information Systems (GIS) are becoming more widely available and interest in their use for the purpose of site selection is increasing. One major advantage in their usage is that planners are no longer tied to geographic units, such as zip codes, that do not adequately represent market areas. One such GIS is described here in regard to the data needed to make it work and how its output can be utilized in planning for retail development. The particular advantage of this system is that it links a GIS to a demographic database that is updated frequently. Population and household forecasts generated from the system are combined with transportation and purchasing information to generate demand forecasts for retail and restaurant space.

Introduction

There have been a number of theoretical and methodological advances in site selection/location studies over the last two decades. As a result better models and more useful applications are evident in the literature. On the theoretical side, as Brown (1991:55) has noted, a number of changes have "...conspired to alter the nature and timing of shopping expeditions and to loosen geographic constraints." The impact of less frequent shopping trips, the increase in store size to include a wide range of products, and direct marketing, among other factors, has altered the retail marketing environment.

There have been significant advances on the methodological side, including the development of Geographic Information Systems (GIS)(Goodchild 1991). Improvements in data-based systems, such as the U.S. Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) system, and Claritas' Prizm package have increased researchers' ability to identify market boundaries and determine the numbers and characteristics of persons within those boundaries. Analysts no longer are constrained by traditional geographic units such as census tracts, block groups and zip codes. In many cases, the use of these units, especially the proliferation of zip code analyses, has misrepresented and distorted the characteristics of current and/or potential market areas. By relying solely on existing geographic units, market areas are either over bounded or under bounded because they are defined in terms of aggregates of those units (e.g., census tracts and zip codes). Therefore, the information studied with respect to the defined area distorts our understanding of the "real" market.

The new systems make use of finer geographic detail such as nodes, latitude and longitude, and grid cells to produce highly accurate boundaries. Basically, a researcher is free to identify an existing or potential market area devoid of many of the constraints that arise when geographic units <u>are</u> derived for other purposes are utilized. The area, regardless of shape, is measured or approximated and the GIS aggregates very small units of geography to create the market boundaries.

The major shortcoming in establishing more accurate market boundaries has been that the data needed for the areas identified, for example, population size and the demographic characteristics of those persons, have not been available. However, data vendors and other organizations are now positioned to aggregate micro-data to produce estimates over many units of geography, and these data are now integral parts of the GIS' in place.

This paper focuses on one of the more advanced geographic based data systems in the U.S. and how it can be used to help make retail site selection decisions. Following a brief review of the literature, which addresses how geographic-based data systems interface with existing retail location models, a description and application of the system is presented.

Literature

In their review of retail location process models, Craig, Ghosh and McLafferty (1984:9) note that, "classical central place theory assumes that consumers patronize the nearest outlet," though larger centers of retail activity are found to be particularly attractive. While not without several serious shortcomings, central place theory has provided a useful mechanism for understanding certain types of shopping patterns, and has given guidance in regard to store location decisions. From a geodemographic perspective, the social and economic characteristics of blocks, neighborhoods, tracts or other geographic units are extremely important as long as product and service preferences and shopping behavior can be linked to the characteristics which comprise an actual or potential market area. With respect to a location decision, one simply has to identify the area with the right characteristics and place an outlet within acceptable time/distance parameters. However, not all consumers minimize distance traveled, and the location of other stores, both similar and dissimilar, along with a host of other factors, has an impact on where people shop.

Store choice models (e.g., Arnold, Roth and Tigert 1980; Louviere and Woodworth 1983) focus on the dynamics of how customers make store choices. Distance traveled (central place theory) is but one attribute that can be measured against others such as price, selection of products, and quality of service among competing outlets. Shopping decisions (locations) are made after all factors are considered. Weisbrod, Parcells and Kern (1984) present a model which estimates the impact of individual socioeconomic characteristics, transportation access, and shopping area characteristics on the choice of shopping destinations. Distance traveled (time) is but one of a host of factors considered when a store choice is made. Even in the store choice context, geodemographic based research has an important role to play because unless hypothetical situations are being considered, retail sites are location-based and customers come from identifiable places. These customers have quantifiable demographic and other characteristics.

Goodchild's (1984) ILACS location model makes use of geodemographic data to select optimum retail locations. Rust and Brown (1986) use squared surface density analysis with customer location (residence) data to graph functions for lunch and dinner markets with respect to restaurant location. More recently, Johnson (1989) identifies how ACORN and Pinpoint, two of the main GIS packages available in the UK, can be used to make decisions regarding new store location, site selection, merchandising, and targeting local differences.

In sum, as in the case of store choice models, store location models have a close tie to geodemographic information. Characteristics of real or potential market areas must be known, and the effect of demographic change is clearly an important component of business strategy.

Geographic Reference Systems

The San Diego Association of Governments (SANDAG) maintains a comprehensive information system stored in both nested and micro-area geographic reference systems. Federal census tract boundaries form the basis of the SANDAG's nested geographic system. There are four levels of geography, and the boundaries of one level do not cross those of another. Traffic Analysis Zones, subdivisions of census tracts, are the smallest area of reference. Geographic aggregations of census tracts form the larger subregional areas and major statistical areas.

Independent of the nested system is the micro-geographic reference system. Gridcells form the basis for this system. One gridcell represents a 2000' by 2000' area, roughly 92 acres. There are 16,000 land-based gridcells covering the cordon area.[1] Data are needed for other geographic areas in addition to those specified. The finest geographic unit that will support those needs is the gridcell split by jurisdiction, traffic zone, and over 70 community planning area boundaries. The 25,000 split gridcells refer to the micro geographic reference area (MGRA).[2]

Once again, the advantage of this system over using traditional geographic boundaries is that gridcells, the smallest unit of geography available, aggregated in the appropriate fashion demark clearly and precisely true market boundaries. The use of larger geographic units, particularly zip codes, generates market boundaries that are most often inaccurate and misleading.

GBF/DIME File and ADMATCH

The preparation of micro-geographic area estimates requires the use of the GBF/DIME file and a Geographic Information System (GIS) procedure known as ADMATCH. The development of the DIME system was a result of the Census Bureau's decision to conduct the 1980 Census largely by mail. Its purpose was to code individual addresses to specific geographic areas for tabulation of census data. Because of its importance in SANDAG's regional information system, SANDAG maintains and updates the DIME file continuously. The Census Bureau's TIGER system replaced the DIME file for the 1990 census. However, files equivalent in structure and content to DIME can be extracted from TIGER.

The DIME file is a computerized map that represents both street networks and special features (e.g., a stream bed representing a census tract boundary). Node points occur where a street or special feature intersect another street or feature, come to an end, or change direction. Shape nodes are used to represent curved features. A line drawn between two nodes is a segment. There are 130,000 segments in the San Diego County DIME file. Each segment contains address ranges and geographic area codes (e.g. census tract, block, and jurisdiction) for both sides of the segment, as well as an X-Y coordinate value at each node point.

One of the primary uses of the DIME file is for address matching (ADMATCH). ADMATCH relates an address-level attribute file to street names and address numbers in the DIME file. The purpose of ADMATCH is to attach geographic elements in the DIME file to the address records in the data file. For the applications discussed in this paper, the primary geographic element of interest is the X-Y coordinate.

Population and Housing Estimates

Determining the geographic boundaries for a real or potential market area provides only a portion of the information needed for a location-based business decision. Data on the population living in that area are also required. Estimating population and housing for micro-geographic areas is a two stage process. First, estimates are produced for split census tract areas; census tracts split by jurisdiction boundaries. The split tract estimates control to the California State Department of Finance (DOF) estimates for the county's 19 jurisdictions.[3] Split tract estimates are then disaggregated to the MGRA level of geography.

Split Census Tract Estimates

The split tract estimates use the "Housing Unit Method" which is the most common technique for preparing sub county population and housing estimates (Lowe 1988; Swanson, et al. 1983; Smith and Lewis 1983). This procedure is represented by the equation below:

$$POPEST = (BYHS + HSCHG) * OCCRATE * PPH + GQPOP,$$

where:

POPEST = total population estimate;
BYHS = base year housing stock;
HSCHG = change in housing stock;
OCCRATE = estimated occupancy rate;
PPH = persons per household; and
GQPOP = group quarters population.

Estimates of split tract total housing units reflect changes in housing stock due to new completions, demolitions, conversions, relocations and annexations.[4] These changes update the most current count of housing units. For example, January 1, 1993 estimates use a 1992 base and building permit data for calendar year 1992. This estimation procedure separately accounts for mobile home units based on an annual survey of mobile home parks. Mobile home unit change on individual lots is reported with the building permit data.

Total units multiplied by split tract specific occupancy rates yield the number of occupied units. Apartment vacancy surveys and electric meter data account for changes in occupancy rates since the last census. Estimates of occupied mobile homes come from a survey of mobile home parks. Household population is estimated by multiplying the number of occupied units by the estimated split tract-specific average household size. Separate sizes are used for mobile home and non-mobile home units. Current estimates of household size update census data using a methodology that integrates historical censuses and symptomatic indicators, such as school enrollment, vital statistics counts and changes in housing structure type.

Persons in civilian and military group quarters facilities added to the household population yield a total population estimate. An annual survey of civilian group quarters facilities produces estimates of this population. DOF collects data annually from each military installation about its barracks and shipboard populations.

Master Geographic Reference Area (MGRA) Estimates

The split tract estimates are disaggregated to the MGRA level of geography. These micro-geographic estimates control to the split tract estimates. This two-stage controlling process, split tract to jurisdiction and MGRA to split tract, insures a consistent set of estimates across geographic levels.

Three site-level files are admatched as a first step in preparing the MGRA estimates. Each address receives a X-Y geographic coordinate from the DIME file. Address files of the mobile home parks and group quarters (civilian and military) facilities are derived from lists used to conduct the annual surveys. These files contain the name, address, and activity count (i.e., total and occupied mobile homes or group quarters population) for each site.

Non-mobile home housing units are estimated from the San Diego Gas and Electric Company (SDG&E) customer service information. This file contains the address and the number of dwelling units associated with each residential electric meter. Master-metered apartments have more than one dwelling unit serviced by a meter. This file includes all meters, whether off or on. These data represent total housing stock, including vacant and seasonal units.

The initial admatch of the SDG&E data involved

TABLE 1
ILLUSTRATION OF THE POINT-TO-POLYGON ACTIVITY ASSIGNMENT PROCEDURE

Address	Address Coordinate X	Y	Activity Type	Activity Count	MGRA ID
2142 Diamond	1156800	635200	SF unit	1 unit	6
1921 Diamond	1156200	635100	Apartment	100 units	6
Total MGRA				**101 units**	**6**
756 Santa Rose	1963000	752000	MH park	30 units	1052
				2 vacant	1052
1022 Freda LN	1825150	792100	Dormitory	200 people	1620
502 Wright	1763592	826500	Mil. GQ	1575 people	9827

roughly 800,000 meters as of January 1, 1986. SANDAG receives annual updates of new meters added to the SDG&E system. These incremental files are admatched and appended to the master file bench marked to January 1, 1986. For example, January 1, 1990 estimates were based on the 1986 meters plus new meters added between January 1, 1986, and December 31, 1989 (127,000 meters). This incremental update avoids the time and cost of admatching each address every year.[5]

The X-Y coordinate assigned during the admatch is the basis for allocating total and occupied mobile homes, group quarters population, and non-mobile home housing units to their appropriate MGRA. This allocation uses an GIS procedure known as POINT-TO-POLYGON INTERSECTION (PIP). A series of X-Y coordinates defines the boundaries of each MGRA polygon and the MGRAs are assigned a unique ID. For each address, PIP compares the site's X-Y coordinate with the MGRA polygon boundary coordinate and determines the MGRA location for the activity. After the PIP, the data are aggregated for each MGRA. Table 1 illustrates the PIP procedure and resultant MGRA information.

The housing unit method is then applied to the MGRAs. The PIP produces estimates of: 1) total housing stock, less mobile homes; 2) total and occupied mobile homes; and 3) civilian and military group quarters population. Applying split tract occupancy rates to the total housing stock yields an estimate of occupied units less mobile homes. Every MGRA within a split tract receives the same occupancy rate. Estimates of household population by MGRA also derive from split tract population per household rates for mobile homes and non-mobile home units.

As mentioned earlier, the MGRA estimates control to the split tract estimates. For mobile homes and group quarters population the control factors are essentially 1.00. The split tract and MGRA estimates both use the same data source (i.e., annual surveys). The MGRA total housing stock estimate requires more extensive adjustment before applying the occupancy and person per household rates.

There are upward and downward biases in the utility address data. SDG&E has no mechanism for reporting meters lost to demolition or fire or distinguishing between meters for new units and those that replace meters on existing units. Downward bias occurs because not all of the residential meters can be successfully matched to a DIME link. This happens for two main reasons: (1) missing streets or addresses in the DIME file, and (2) incompatible address between the meter data and DIME file (e.g., P.O. Box). Missing streets and addresses are the most serious problem. A current DIME file is essential to the accuracy of the MGRA estimates, especially in fast growing areas.

Neither bias caused a serious problem in the 1990 MGRA estimates. The average housing unit adjustment factor was five percent across all split tracts. MGRA units had to be adjusted by more than 10 percent in only 15 split tracts. This relatively minor degree of adjustment shows that the MGRA estimates were very consistent with the split tract housing counts. This consistency is important since excess controlling would distort the accuracy of the MGRA housing counts.

Creating Estimates for Unique Geographic Areas

Once the MGRA population and housing data base is created, estimates for unique geographic areas are developed by aggregating MGRA information. Several GIS techniques are used to define a geographic area according to MGRA boundaries. For example, estimates are desired for a market area or a 3 mile radius from a given intersection. The boundaries for these areas are first digitized and assigned X-Y coordinates (i.e., creating a polygon). Using a procedure known as POLYGON-TO-POLYGON OVERLAY, the geographic polygon boundary is computer matched to the MGRA polygon boundary file. The geographic area is defined by the MGRAs whose centroids (most central X-Y coordinate) fall within the geographic area's polygon boundary.

A procedure known in ARC/INFO as NETWORK is used to determine estimates for, say, an area 10 minutes from a given site. NETWORK will identify the centroids of MGRAs within 10 minutes of a site, based on the street pattern and street type that connect the MGRAs. For example, MGRAs connected by freeways are closer to each other in terms of time compared to MGRAs that are only linked by local roads.

It is important to note that MGRA boundaries yield an approximate definition of a geographic boundary. If a polygon boundary splits an MGRA, then the MGRA is only allocated to the area if its centroid falls within the polygon. All information contained in a border MGRA, if assigned, is attributed to the estimate. This all or nothing assignment of MGRAs does not seriously distort the estimate, primarily because of the small geographic size of an MGRA. Estimates where MGRA information was split based on the fraction of the MGRA inside the polygon were compared to estimates based the all or nothing assignment. The two estimates were not different enough, in most comparisons, to warrant the extra time and cost involved in implementing a split MGRA approximation routine.

There are instances where the MGRA approximation is not precise enough. The obvious situation is where the geographic area is much smaller than an MGRA, although this rarely occurs.[6] More common is the circumstance where even a small error could lead to an erroneous decision, such as the activity under an airport influence area. For these instances, the PIP procedure places the housing units and group quarters population directly into the area defined by the polygon, bypassing the MGRA data set. Split tract level vacancy rates and household sizes are used to complete the estimate. This allows a very precise accounting of the current housing unit activity within the geographic area, but is considerably more costly than approximating area using the MGRAs.

One question that might be raised is how accurate are these estimates. A recent analysis compared this GIS's aggregated estimates with 1990 census counts for census tracts and blocks. The average absolute percent errors were 8% and 13% respectively for tracts and blocks, making these estimates as least as reliable any estimates produced for this level of geographic detail. Moreover, San Diego county is like many other markets where population change, in this case growth, is very rapid. Without current population and household estimates for the post-1990 period, planners would be forced to rely on 1990 census counts which have become outdated quite quickly. In addition, demographic forecasts produced for the market areas of interest extend the planning horizon greatly.

Application

La Mesa is a city of 52,000 persons located 15 miles east of downtown San Diego. This city grew 5.2 percent during the 1980s, well below the 34 percent growth rate of the region. A substantial amount of La Mesa's growth occurred through redevelopment activities primarily because of its limited supply of vacant developable land. La Mesa contracted with SANDAG to conduct a market analysis for a major redevelopment area known as Alvarado Creek. The overall objective of this study was to forecast Alvarado Creek's market demand for additional commercial retail space, office space and hotel/motel accommodations.

The boundaries of the market area were determined by a number of factors, including the type of development, accessibility, and location of competing facilities. Our client was primarily interested in primary and secondary market areas defined in terms of travel times from the site. The primary market area was within a 15 minute drive time from the redevelopment site, while the secondary market area encompassed drive times from 15 to 30 minutes. Current and future population and households for the primary and secondary market areas were used to help determine Alvarado Creek's development potential. Table 2 shows estimates of the 1988 population and households data along with their 2010 forecasts.[7] These figures were derived by accumulating information from the MGRAs that defined each market area.

There were 268,800 households and 692,200 persons in the total market area in 1988, 30 percent of the

TABLE 2
ALVARADO CREEK MARKET AREA
POPULATION AND HOUSEHOLDS, 1988-2010

	1988	2010	1988-2010 Change Number	Percent
Primary Market Area				
Population	116,900	123,800	6,900	5.9%
Households	44,900	49,200	4,300	9.6%
Secondary Market Area				
Population	575,300	672,400	97,100	16.9%
Households	223,900	274,900	51,000	22.8%
Total Market Area				
Population	696,200	796,200	104,000	15.0%
Households	268,800	324,100	55,300	20.6%

county's residential activity. A 55,300 increase in households is projected to occur over the next 22 years. This represents a 21 percent increase, less than one-half of the change forecast for San Diego county. By 2010, the Alvarado Creek market area will comprise 25 percent of the region's residential activity, a decline in proportional representation from now.

The primary market area is forecast to grow more slowly than the secondary market area. An additional 4,300 households and 6,900 persons are forecast for the primary market area by 2010. As stated earlier, the main reason for the slow rate of future growth is the lack of developable land within the primary trade area. Households and population in the secondary market area will grow by 23 percent and 17 percent, respectively. An important factor which contributes to faster growth in the secondary market area is the development of a general aviation airport into an industrial park. Manufacturing employment within the secondary market area is forecasted to increase by 11,800 jobs over the next 22 years. A majority of these new jobs (85 percent) will locate in the airport industrial park.

The Alvarado Creek redevelopment area can anticipate growth in potential clientele. However, this redevelopment area loses its share of the regional market place. Plans and strategies are needed to increase its attractiveness to shoppers in the secondary market area and beyond. Two major transportation improvements will increase the accessibility of the redevelopment area. One is the extension of light rail service into this area, and the other is the connection of two freeways near Alvarado Creek. These transportation improvements should help expand the secondary market eastward into places that will grow faster than the county over the next 22 years.

Market Demand

Characteristics of the market area's population and housing are important, but they are not the only factors that determine market potential and the prospects of a proposed retail site. Other factors such as expenditure patterns, buying power and existing competition should be part of the information matrix. The multi-use nature of the Alvarado Creek development required information about current and future demand for retail space and restaurant facilities.[8]

Retail Demand

Consumer demand available to support the proposed Alvarado Creek retail development was estimated using the number and buying power of households within the market area. The number of households have already been presented in Table 2, while their buying power was based on expenditure patterns and income levels. The annual household expenditure pattern for retail goods and services was based on survey data for households within San Diego county (U.S. Bureau of Labor Statistics 1980: table 150). The household expenditure pattern was applied to 1980 income figures for the Alvarado Creek area to derive average household expenditures for retail goods and services in the primary and secondary market areas. These averages were adjusted to reflect the rise in consumer prices and changes in household income distributions from 1980 to 1988 (SourcePoint 1990).

Several other adjustments were needed to determine if the level of household expenditures could support additional retail space. The expenditure averages represent all consumer purchases. However, Alvarado Creek is planned as a community shopping center and will not carry the full variety of goods and services of a regional shopping center. The expenditure averages were lowered by 25 percent to reflect the fact that consumers cannot meet all of their retail needs from a community shopping center. These adjusted averages applied to the 1988 households yielded total expenditures of $656 million and $3.0 billion dollars for the primary and secondary market areas, respectively.

A second adjustment reflects the proportion of total sales that comes from each market area. Household spending patterns in urban areas suggest that roughly 80 percent of household expenditures made at community shopping centers come from the primary market area and 20 percent from the secondary market area. Applying these capture rates against total expenditures resulted in total expenditures in 1988 of $525 million for the primary market area and $598 million for the secondary area. The last adjustment takes into account existing competition. Sales in existing retail sites was estimated from their square footage using the median annual sales volume per square foot of gross leasable area (Urban Land Institute 1987: tables 3-12, 5-12, 6-12). The medians varied according to the type of retail site (i.e., strip commercial or neighborhood, community or regional shopping center).

For the secondary market area, existing sales exceeded the $598 million in household expenditures for retail goods in that area; therefore, no additional retail space would be supported by these households. Expenditures did exceed sales in competing sites by 31 million dollars in the primary market area. Based on the median sales value per square foot for community shopping centers ($172), this translates into approximately 182,000 square feet of additional retail space that could be supported by existing (1988) households within the Alvarado Creek market area. Assuming the same base year factors and adjustments for future prices and income level, Alvarado Creek would support only an additional 11,000 square feet of retail space between 1988 and 2010. This is due to the relatively slow growth in future households in the primary market area.

Restaurant Demand

Demand for additional restaurants was calculated in the same manner as described above for all retail space. The expenditures and other factors were specific to restaurants rather than for all retail uses. Only households in the primary market area had sufficient expenditures to support additional restaurant facilities in 1988. The household expenditures were $89.3 million compared to $86.8 million in sales at existing restaurants. Sales at existing restaurants was based on California State Board of Equalization information on retail sales for restaurants.

The difference of $2.5 million between household expenditures and restaurant sales represents net potential expenditures that could be captured in 1988 by constructing new restaurants at Alvarado Creek. This translates into an additional 21,400 square feet based on the sales per square foot ($117) for restaurants located in community shopping centers. Restaurants in community shopping areas generally range in size from 1,800 to 3,600 square feet. Using the midpoint of this range would imply that the market area could support eight additional restaurant sites in 1988.

The expected growth in households will enable Alvarado Creek's primary market area to support additional restaurant space in the future. Approximately 190 new housing units are needed to support one average size restaurant in the market area. This figure was derived by dividing the retail sales for all restaurants ($77.1 million) by the number of restaurants in the market area (217), then dividing that result by the restaurant expenditures for a household ($1,900). Growth of 4,300 households in Alvarado Creek's primary market area would support an additional 23 restaurants (approximately 62,000 sq. ft.) between 1988 and 2010.

Once these forecasts are generated, macro-level (from the perspective of a developer) and micro-level (from the perspective of an individual store owner) planning can proceed. From the macro perspective, the issues of importance concern the location of multiple retail and service sites in order to maximize overall economic activity, and thus return on investment. From a micro perspective, the major focus is on potential sales in the context of a matrix of competitor or supportive businesses.

Conclusion

This paper describes procedures for developing accurate population and housing estimates for micro-geographic areas. The methodology shows the power and versatility that GIS technology linked to reliable and recent data bases offer for data retrieval. Micro-geographic areas provide the basis for aggregating demographic information that accurately reflects the boundaries for virtually any user defined area. The procedures described here are not the only way to deliver demographic profiles for custom areas. They do, however, overcome limitations found in other data retrieval procedures. These include the: (1) inability to reflect major population changes; and (2) tendency to view development as evenly distributed across a geographic area.

The method's value was demonstrated in an application that involved forecasting of the potential demand for commercial services at a proposed redevelopment site. While showing the utility and flexibility of GIS data retrieval techniques the paper also points out that factors in addition to population and housing demand must be considered when evaluating potential retail sites.

Currently, the data retrieval methods discussed in this paper are becoming more widespread. Most large urban areas have access to the data sources and technology described. A primary reason for the lag in the use of GIS data retrieval systems is the lack of accessible computerized street files, comprehensive address lists and GIS software. However, the single biggest event that will lead to a more widespread dissemination of GIS applications is the development of the TIGER file. For the first time, there is a comprehensive computerized street file for the United States.

TIGER makes computerized street information readily available and there are a number of vendors who are currently exploiting its commercial value. Address lists are now available from a variety of private sources. Utility companies represent another potential source of address lists. Finally, the proliferation of GIS software will be central to information systems development in the 1990s. These tools are now becoming available to a wide range of users which should only increase their application over time.

TIGER does have it faults and there is the issue of how TIGER will be updated to reflect current development. Moreover, the site selection use of TIGER is only as good as the data available to describe and analyze the market area identified. As the data from the 1990 census become outdated, systems such as that described here become even more valuable because they provide new information about the population living in the area.

Suggestions for Future Research

Improvements in data availability and advancements in the hardware and software needed for more sophisticated GIS will enable planners to utilize even more complex analyses in their work. The key to improved applications lies in the ability of marketers to bring together diverse datasets so that all of the relevant information needed for problem solving is incorporated into these systems.

The authors wish to thank Jackie Lynch for her assistance in the preparation of this manuscript.

*****Footnotes*****

1. This micro-geographic data base pertains to the western half of San Diego County. In 1990, the cordon area contained over 99 percent of the county's population and housing units. The micro-geographic reference area was recently expanded to cover the entire county. It is based on 1990 census blocks split by several geographic boundaries. This MGRA contains 25,915 pieces of geography.
2. MGRA boundaries are computerized using California State Plane X-Y coordinates. These digitized files are manipulated with the ARC/INFO Geographic Information System software. An excellent overview of the GIS procedures used in this paper is found in Environmental Systems Research Institute (1989).
3. A detailed discussion of the methodology for producing split-tract estimates is found in SANDAG (1988a) and Tayman (1986).
4. SANDAG receives building permit activity by census tract from the 19 jurisdictions in San Diego County.
5. The 1986 SDG&E data had a match rate of 95 percent (760,000 out of 800,000 addresses). MGRA estimates for the 1990s will be based on a complete admatch of meters as of April 1, 1990 and incremental meter address updates for each year after.
6. An area of 350 acres (4 MGRAs) is the smallest area for which minimally acceptable estimates can be produced.
7. Long-range population and forecasts by MGRA are based on land use models which are described in SANDAG (1988b). Putman (1983) contains an excellent overview of the development and use of spatial interaction land use models for small area forecasting.
8. The client also requested demand forecasts for office space and hotel rooms. This information is not presented since the focus of this paper is on retail activities. The results of the analysis of office space and hotel rooms is available from the authors.

*****References*****

1. Arnold, S. J., V. Roth, and D. Tigert, "Conditional Logit MDA in the Prediction of Store Choice," in A. A. Mitchell (ed.), *Advances in Consumer Research* IX, Ann Arbor, MI: Association for Consumer Research, pp. 665-670, 1984.
2. Brown, Stephen, "Retail Location Theory: Retrospect and Prospect," *Irish Marketing Review*, Vol. 5, No. 2, pp. 52-59, 1991.
3. Craig, Craig S., Avijit Ghosh and Sara McLafferty, "Models of Retail Location Process: A Review," *Journal of Retailing*, Vol. 60, No. 1, pp. 5-36, 1984.
4. Environmental Systems Research Institute, *ARC INFO Users Guide*, Vol. I. Redlands, CA, 1989.
5. Goodchild, Michael F., "Geographic Information Systems," *Journal of Retailing*, Vol. 67, No. 1, pp. 3-15, 1991.
6. Goodchild, Michael F., "ILACS: A Location Allocation Model for Retail Site Selection," *Journal of Retailing*, Vol. 60, No. 1, pp. 84-100, 1984.
7. Johnson, Maureen, "The Application of Geode-

mographics to Retailing--Meeting the Needs of the Catchment," *Journal of the Market Research Society*, Vol. 31, No. 1, pp. 7-36, 1989.
8. Lowe, Theresa J., "A Resurrection: The Potential of Postal Survey Data in Improving Housing Unit Population Estimates." Paper presented at the annual meeting of the Population Association of America. New Orleans, Louisiana, 1988.
9. Louviere, Jordan J. and G. Woodworth, "Design and Analysis of Simulated Consumer Choice of Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, Vol. 20, pp. 350-367, 1983.
10. Putman, Stephen H., *Integrated Urban Models*, London, Pion, 1983.
11. Rust, Robert T. and Julie A. Brown, "Estimation and Comparison of Market Area Densities," *Journal of Retailing*, Vol. 62, No. 4, pp. 410-430, 1986.
12. San Diego Association of Governments, *Population and Housing Estimation System: Revised Users Manual and Technical Documentation*. San Diego, CA, 1988a.
13. San Diego Association of Governments, *Series 7 Regional Growth Forecast*, Vol. II: Technical Users Manual for PLUM and SOAP. San Diego, CA, 1988b.
14. Smith, Stan and Bart Lewis, "Some New Techniques for Applying the Housing Unit Method of Local Population Estimation: Some Further Evidence," *Demography*, Vol. 17, pp. 323-339, 1983.
15. SourcePoint, *1988 Household Income Estimates*. San Diego, CA, 1990.
16. Swanson, David, Barbara Baker, and Jo Van Patten, "Municipal Population Estimation: Conceptual and Practical Features." Paper presented at the annual meeting of the Population Association of America. Pittsburgh, Pennsylvania, 1983.
17. Tayman, Jeff, "Population and Housing Estimation system for Census Tracts." Paper presented at the annual meeting of the Population Association of America. San Francisco, California, 1986.
18. Bureau of Labor Statistics, *Handbook of Labor Statistics*, U.S. Government Printing Office, Washington, D.C., 1980.
19. Urban Land Institute, *Dollars and Cents of Shopping Centers*. Washington, D.C., 1987.
20. Weisbrod, Glen E., Robert Parcells, and Clifford Kern, "A Disaggregate Model for Predicting Shopping Area Market Attraction," *Journal of Retailing*, Vol. 60, No. 1, pp. 65-83, 1984.