Applications of Geographic Information Systems for the Analysis of Apartment Rents

*Douglas S. Bible**
*Cheng-Ho Hsieh**

Abstract. **This study is the first to incorporate comprehensive regional factors into the analysis of the variations of apartment rent in a particular metropolitan area. A Geographic Information Systems (GIS) procedure is developed to generate regional variables for the analysis. Results show that not only the individual apartment's characteristics, but also the regional factors are important in determining apartment rents.**

Introduction

Hedonic pricing models have been used to study the effects of housing markets, especially the effects of various attributes upon apartment rents in recent years. Sirmans and Sirmans (1992) and Benjamin and Lusht (1993) found significant effects of property management upon rents, while Sirmans, Sirmans and Benjamin (1989) observe the effects of amenities and external factors on rents. Although the later study does consider the number of major employees in a locational sector as a proxy for commuting time to work, these models in general do not consider the effects of regional factors and neighborhood characteristics on apartment rents.

Graham and Bible (1992) note that neighborhood and distance factors, as well as physical attributes, are important in the classification of apartments. No studies are found that use GIS techniques to derive regional variables (such as distance to work or average income for a uniquely defined neighborhood) useful for studying the apartment market.

Aalberts and Bible (1992) use GIS analytical techniques to assist in the development of a mortgage foreclosure model. Their spatial analysis suggests a potential relationship between income, race and distance variables, and mortgage foreclosure rates. GIS tools are used to calculate those variables. However, the results find no significant relationships.

Thrall, Sidman, Thrall, and Fik (1993) demonstrate the usefulness of GIS software to study large spatially related datasets. Their study uses census tract housing data and GIS techniques to forecast single-family residential development for an entire country.

In examining the functional requirements of Geographic Information Systems for real estate research and analysis, Thrall and Marks (1992) observe that a Geographic Information System (GIS) is a powerful tool that improves research productivity by allowing calculation of measurement and modification of data that would otherwise be difficult or impossible to execute. The increased productivity allows and encourages researchers and practitioners to undertake analysis not previously attempted. They suggest that a hedonic model of apartment building rents could include traditional

*Department of Economics and Finance, Louisiana State University-Shreveport, Shreveport, Louisiana 71115.
Date Revised—August 1994; Accepted—December 1994.

amenity and design measures, neighborhood characteristics, and accessibility measures with the latter being generated as output from GIS. This approach is followed in the current study.

In this paper, a study of apartment rents in the Shreveport-Bossier (Louisiana) Metropolitan Statistical Area is undertaken using GIS to emphasize the spatial aspects of the local marketplace. A traditional hedonic model that includes GIS-generated regional/neighborhood variables, variables often omitted in previous studies, is formulated to explain apartment rents. Marks, Stanley and Thrall (1994) evaluated several GIS software programs, including Atlas GIS, based on functionality, ease of use and documentation. They noted that Atlas performed most of expected GIS functions, but was weak in statistical functionality. In the current study, Atlas[1] GIS software was used to perform numerous geographic or spatial analysis. The mapping capabilities and the ability to merge databases (for both points and regions) with geographic information (such as TIGER Files), common for most GIS software, were critical for the research study.

Several advantages of GIS applications which Thrall and Marks discuss are used in the current paper. These include: 1) GIS technology makes it possible to measure and evaluate absolute location. Features, such as apartment buildings, are precisely located according to latitude and longitude coordinates. This allows visualization of real estate data in a map form that includes various landscape features that have been digitized by the U.S. Bureau of the Census and commercial vendors. The base map for this study includes the TIGER/Line data files with the names of roads and the range of addresses that occur along each road segment; 2) GIS functions are distinguished from relational database management capabilities through the software's ability to perform spatial measurements and perform polygon operations. In this study, distance from each apartment building to universities and to major shopping centers is measured using batch mode processing. These distances are retained by the GIS software for application in an econometric analysis; 3) "GIS polygon operations manipulate attribute data assigned to the polygon or data that otherwise fall within the bounds of the polygon. Polygon operations allow the user to add new polygons" (Thrall and Marks, 1993, p. 56). The current study uses both census tracts and user-defined circular regions (with a one-mile radius) to derive regional variables for each apartment.[2]

The data and the methodology are described in section two; section three explains the GIS procedures for generating regional/neighborhood variables. Section four discusses the empirical results, and the fifth section concludes this study.

Data and Methodology

The database for this study consists of the nonsubsidized apartment complexes located in the Shreveport-Bossier metropolitan area. The current rents, apartment characteristics data, and regional variables generated by GIS are collected for all of the eighty-one apartment complexes.

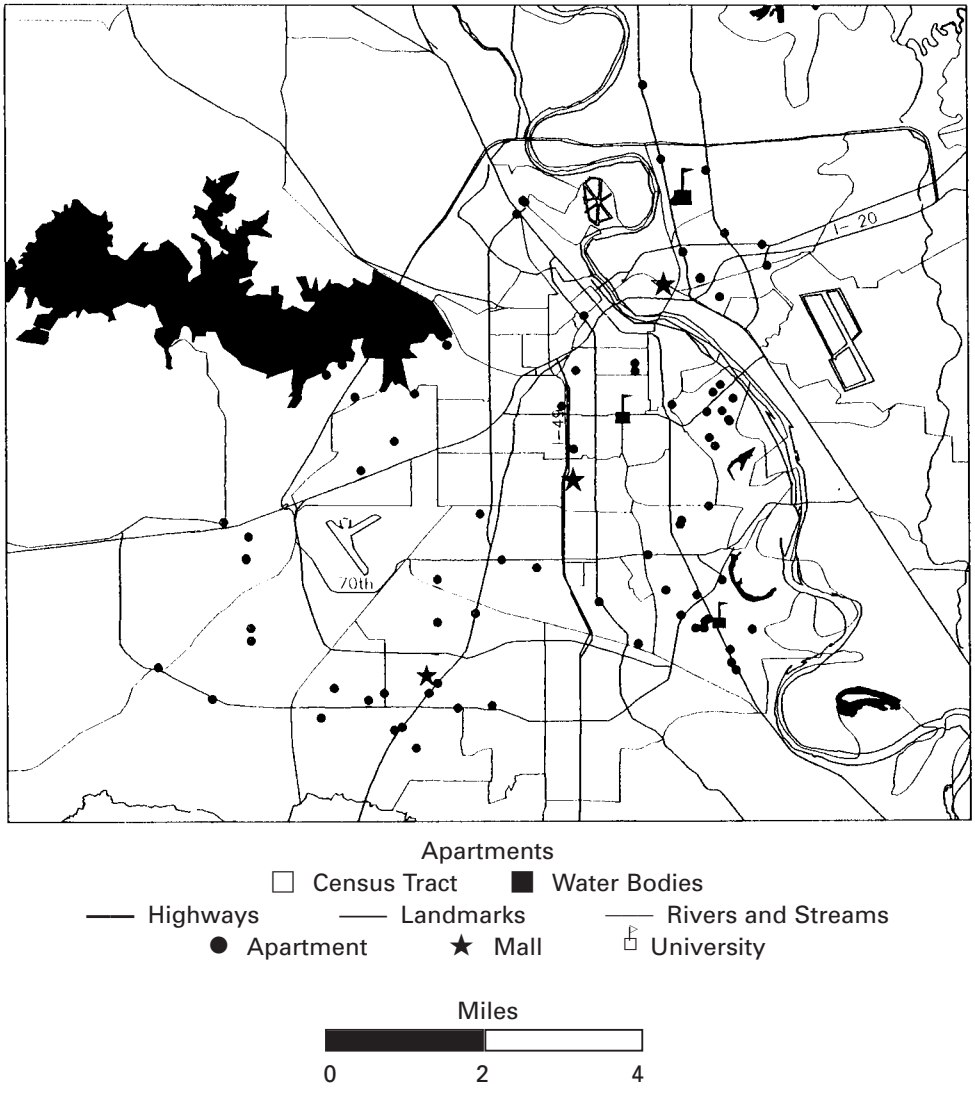
Apartment characteristics data include the year the apartment was built; the number of bedrooms and square feet for each unit; the existence of a pool, fireplace and tennis courts; and the total number of units of the apartment. From these data, five variables are generated to be used as the independent variables for the following basic model:

$$RENTSQFT = a_0 + a_1 AGE + a_2 POOL + a_3 TENNIS + a_4 FIREPLACE + a_5 \mathrm{L}(SIZE) + e \ ,$$

where *RENTSQFT* is the weighted average rent per square foot for the apartment complex; *AGE* is 1993 minus the year the apartment was built; *POOL* is a dummy variable that takes on a value of one if the apartment has a swimming pool and zero otherwise; *TENNIS* is a dummy variable that takes on a value of one if the apartment has a tennis court and zero otherwise; *FIREPLACE* is also a dummy variable that takes on a value of one if the apartment has a fireplace and zero otherwise; and L(*SIZE*) is the natural logarithm of total number of bedrooms of the apartment complex.

Exhibit 1, generated with the GIS program, shows the location of apartment buildings along with the major roads and the census tract boundaries, as well as the location of

Exhibit 1
Shreveport–Bossier Apartments



colleges and shopping centers. To develop neighborhood characteristics for the apartments in this study, a separate layer consisting of regions (circles with a one-mile radius around each apartment building) were established with the GIS software.

After the apartment regions are established, several regional variables are generated using GIS. These regional variables may affect the desirability of apartments in a fashion similar to the residential marketplace. For example, Bible and Grablowsky (1984) find that multifamily residential units located in neighborhoods with special restorative zoning controls increase in value at a significantly higher annual rate than those with no special restorative codes during the same time period. Bible and Crunkleton (1983) also find location (based upon quality of schools, accessibility to employment and shopping, and the general physical appearance of the neighborhood) to be important in explaining apartment value.

Using GIS, virtually an unlimited number of regional variables can be generated. It would be inappropriate to attempt to incorporate all possible variables into the model. Eleven regional variables are generated and are added into the basic model for an ordinary least squares regression analysis. However, the collinearity diagnostics generated by SAS PROC REG show that a severe multicollinearity problem exists. To reduce the effects of this problem, the eleven variables are grouped into five sets based on common characteristics of variables. The basic model is then extended by one of the five sets of regional variables. Thus a total of six models, including the basic model, are formulated.

Since no theoretical guidance exists as to how regional factors affect apartment rents, this study does not attempt to specify the "correct" model. Rather, the emphasis is on how GIS techniques can be used to discover the potential importance of regional factors in determining apartment rents. The five sets are described as follows:

- Occupation Variables

  PMGMT: number of managers in the region divided by total number of workers in the region;
  PSVWK: number of service workers in the region divided by total number of workers in the region;
  PFARM: number of farmers in the region divided by total number of workers in the region;

- Income Variables

  L(*AVGINC*): the natural logarithm of average household income;
  PRTINC: residential rental expense as a percentage of household income;

- Competition Variables

  AGEDA: the difference between the age of the apartment and the average age of other apartments in the region;
  NAPT: total number of other apartments in the region;

- Time Variable

  TWK: average time to get to workplace from each region;

- Distance Variables

SCHOOL: the distance to the closest of three colleges in the Shreveport/Bossier area;

SHOP: the distance to the closest of the three major shopping malls in the Shreveport/Bossier area;

WORK: the distance to the closest of the four major business centers[3] in the Shreveport/Bossier area.

The expected outcomes for the variables in the model are now discussed. Overall, it is anticipated that higher incomes, greater socioeconomic status (represented by occupation), newer buildings, nearness to work, schools and shopping will contribute on a positive basis to the rent levels for the apartments. With the occupation variables, it is anticipated that a higher percent of managers and service workers, relative to lower income blue collar workers, would lead to a higher rent. However, the percentage of farmers may have little if any effect.

Household income is expected to have a positive impact on apartment rent, but the percentage of income spent on rent may have a negative effect since lower income households tend to spend a greater percentage of their total income on housing (Sullivan, 1990). The age variable is expected to be negative since older apartments generally show more depreciation and thus rent for less. The number of other apartments in the region may represent the local competition and is expected to have a negative influence on rents.

The time variable, *TWK*, is the average time for people living in the region to get to their work place. It is assumed that average time to get to the work place for people living in the apartments is equal to *TWK*. Therefore *TWK* is expected to have a negative effect on apartment rent.

The distance variables are expected to be negative since being close to schools and work are viewed as being favorable housing attributes. Being near to shopping is also expected to have a positive effect on rent, however, negative effects may also be evident due to increased traffic and noise.

The value for these regional variables can be calculated through the use of the GIS software. Details regarding the use of GIS tools are described next.

The GIS Procedure

Geographic files, attribute files and datapoint files are used in this study. Mapping and spatial analysis generally requires some type of base map for the geographic area under consideration. The Caddo and Bossier Parish Louisiana TIGER map files, are used for this study.[4]

The individual apartment data, including the locational information as well as the apartment characteristics data discussed in section two, are imported from a spreadsheet into the GIS as a datapoint file. The apartments are then assigned specific geographic locations with latitude and longitude fields through the use of an address matching procedure.

The attribute files contain housing and population census information. This census information, available on CD ROM disks, is downloaded to a spreadsheet, imported into the GIS, and merged with the geographic files according to census tract. The

apartment regions developed by the authors are added as a layer onto the geographic file. This allows for the development of numerous regional variables based on the census population and housing data. The GIS program assigned attribute values of the underlying census tracts to the apartment regions based upon a geographic area weighted procedure that assigns values according to the proportion of each tract that is contained in the apartment region. For example, if one half of the area of a census tract was included within an apartment region, then one half of the value of the population and housing variables would be included for that region. Values such as number of homes would be summed, while averages such as family income, would use a weighted average.

Results

Simple correlation analysis generated by SAS PROC CORR[5] shows that *AGE, PRTINC, AGEDA, SCHOOL,* and *TWK* are negatively and significantly correlated with *RENTSQFT*; while *POOL, FIREPLACE* and *PMGMT* are positively and significantly correlated with *RENTSQFT.* Exhibit 2 provides the summary statistics for all the variables. Exhibit 3 shows the regression results of the six models using *RENTSQFT* as the dependent variable. Model 1 is the basic model that includes only the five apartment attribute variables. Models 2 through 6 represent the basic model extended by one of the five sets of regional variables. *AGE* shows negative and significant effects in three of the seven models, indicating that the higher the apartment age, the lower the rent per square foot. *POOL* and *FIREPLACE* are positive and significant, indicating that the existence of a swimming pool and/or a fireplace increases the rent. *TENNIS* is not significant. L(*SIZE*) is negative and significant, indicating that rent per square foot reflects possible economies of scale in apartment operations.

Model 2 includes the occupation variables as the regional variables. Civilian occupations are categorized into four groups: management, service, farmer, and blue collar. *PMGMT* is the number of management workers divided by the total; *PSVWK* is the number of service workers divided by the total, and *PFARM* is the number of farmers divided by the total. Thus blue collar workers are the omitted class. Results show that, as expected, *PMGMT* and *PSVWK* are positive and significant, indicating that the higher the percentage of management and/or service workers (a measure of socioeconomic status), relative to blue collar workers in the region, the higher the apartment rent. That the variable *PFARM* is not significant is not surprising, since most of the regions had relatively few farm workers.

Model 3 is the basic model plus the income variables as the regional variables. Both L(*AVGINC*) and *PRTINC* are not significant. Model 4 is the basic model plus the competition variables. Both *AGEDA* and *NAPT* are not significant. Surprisingly, the competition in terms of age and number of competitive apartments did not effect the level of rents.

Model 5 is the basic model extended by the time variable as the regional variable. *TWK* is negative and significant, indicating that the longer the time to get to the work place, the lower the apartment rent.

Model 6 includes the distance variables as the regional variables. *SCHOOL* is negative and significant, indicating that the farther the distance from the nearest college, the lower the rent. The significance of this variable reveals that college students constitute an

Exhibit 2
Summary Statistics
(No. of Observations: 81)

Variable	Mean	Std Dev.	Min.	Max.
RENTSQFT	.43	.07	.27	.67
AGE	17.70	8.96	1.00	46.00
POOL	.91	.28	.00	1.00
TENNIS	.41	.49	.00	1.00
FIREPLACE	.44	.50	.00	1.00
SIZE	273.39	135.02	67.00	812.00
PMGMT	.70	.11	.45	.87
PSVWK	.12	.06	.05	.31
PFARM	.01	.00	.00	.03
AVGINC	32711.00	10661.00	12041.00	78834.00
PRTINC	.26	.04	.18	.34
AGEDA	.08	5.63	−12.50	14.33
NAPT	4.53	2.74	1.00	10.00
TWK	17.63	1.97	14.20	23.30
SCHOOL	3.49	2.86	.25	13.34
SHOP	2.74	1.59	.35	9.23
WORK	3.14	1.54	.26	6.24

important part of the apartment rents in this area. *SHOP* is positive but not significant. *WORK* is negative and significant, an indication that the nearness to work is a positive attribute.

Conclusion

This study is the first to incorporate regional factors into the analysis of the cross-sectional variations of apartment rent in a particular metropolitan area. The regional variables are generated using the Geographic Information System (GIS). GIS allowed for the accurate and efficient calculation of distance and regional variables that otherwise would have been difficult and impractical to execute. By doing so, this study also shows several detailed applications of GIS in the context of a commercial real estate database for apartment buildings in a metropolitan area. Individual apartment characteristics and regional factors have been examined and merged through geographic analytical procedures, providing insight into the usefulness of GIS for studying data within a spatial context.

The results show that among the apartment characteristics variables, the age of the apartment shows negative effects on apartment rent. A swimming pool and/or a fireplace command a higher rent. The size variable has a significant and negative effect on apartment rents, indicating the existence of economies of scale on apartment operations.

Of the regional variables generated by GIS, the variables with occupational characteristics such as percentage of management and service workers have positive effects on rent relative to blue collar workers. The amount of time to get to the work place shows negative effects on apartment rent.

The variables with distance measures such as the distance to colleges have a negative effect on rent. However, the distance to shopping malls has no significant effect on apartment rent. The distance to work shows negative and significant effects.

Exhibit 3
Regression Results for Shreveport–Bossier Apartment Market with Dependent Variable Rent per Square Foot

	Model					
Variable	1	2	3	4	5	6
Intercept	.5167[a]	.0895	.8119[b]	.4858[a]	.7033[a]	.5770[a]
	(5.894)	(.579)	(2.528)	(5.370)	(6.531)	(6.361)
AGE	−.0015	−.0025[b]	−.0015	−.0005	−.0019[b]	−.0025[b]
	(−1.510)	(−2.530)	(−1.476)	(−.411)	(−1.992)	(−2.495)
POOL	.1024[a]	.0850[a]	.0987[a]	.1038[a]	.0883[a]	.0910[a]
	(3.752)	(3.062)	(3.461)	(3.679)	(3.313)	(3.352)
TENNIS	−.0038	.0065	−.0083	−.0051	.0015	−.0012
	(−.200)	(.356)	(−.423)	(−.270)	(.085)	(−.063)
FIREPLACE	.0303[c]	.0243	.0283	.0298[c]	.0300[c]	.0286[c]
	(1.701)	(1.445)	(1.571)	(1.674)	(1.761)	(1.675)
L(*SIZE*)	−.0311[c]	−.0286[c]	−.0271[c]	−.0294[c]	−.0281	−.0275[c]
	(−1.944)	(−1.889)	(−1.641)	(−1.831)	(−1.828)	(−1.778)
PMGMT		.4980[a]				
		(3.413)				
PSVWK		.9178[a]				
		(3.045)				
PFARM		−1.3692				
		(−.704)				
L(*AVGINC*)			−.0236			
			(−.847)			
PRTINC			−.2563			
			(−1.000)			
AGEDA				−.0024		
				(−1.398)		
NAPT				.0009		
				(.317)		
TWK					−.0105[a]	
					(−2.770)	
SCHOOL						−.0095[a]
						(−3.187)
SHOP						.0088
						(1.425)
WORK						−.0135[b]
						(−2.043)
R^2	.2733	.3802	.2848	.2934	.3416	.3667
F-Value	5.642	5.520	4.152	4.330	6.399	5.211

t-statistics in parentheses
[a]significant at 99% level; [b]significant at 95% level; [c]significant at 90% level

Variables proxy for competition in the region such as the difference between the age of the apartment and the average age of other apartments in the region and the total number of other apartments in the region have no effect on apartment rent. Interestingly, variables with wealth characteristics such as the average household income in the region show no effect on apartment rent.

These results suggest that regional factors are potentially important in determining apartment rents, in other words, location matters. In this study we develop a GIS procedure to generate regional variables. This procedure can be used for other graphic analyses; for example, potential market areas may be mapped and used to generate regional variables for market analysis. In addition, this procedure can be used for the study of site selection for commercial real estate uses. For example, a retailer may require a minimum population and income base within a certain distance of a new site location; GIS procedures could be used to develop regions that meet the specified criteria.

Notes

[1]Atlas is a Desktop GIS product of Strategic Mapping Inc., Santa Clara, California.
[2]GIS software allows the user to define regions based on a circle of any desired radius.
[3]They include General Motors, AT&T, Barksdale Air Force Base, and Central Business District.
[4]The TIGER (Topologically Integrated Geographic Encoding and Referencing) system is a digital (computer-readable) map database that contains data for census geographic features (such as roads, railroads and rivers); feature names and classification codes; Federal Information Processing Standard codes for census tracts, blocks, cities, and townships; and within metropolitan areas, address ranges and ZIP codes for streets.
[5]This software is published by SAS Institute, Inc., SAS Circle, Box 8000, Cary, North Carolina 27512-8000.

References

Aalberts, R. J. and D. S. Bible, Geographic Information Systems: Applications for the Study of Real Estate, *Appraisal Journal,* 1992, 60:4, 483–92.

Benjamin, J. D. and K. M. Lusht, Search Costs and Apartment Rents, *Journal of Real Estate Finance and Economics,* 1993, 6:2, 189–97.

Bible, D. S. and J. R. Crunkleton, The Effects of Financing on the Sale of Multi-Family Properties, *Real Estate Appraiser and Analyst,* 1983, 49:2, 33–37.

Bible, D. S. and B. J. Grablowsky, Restorative Zoning Effects on the Valuation of Multi-Family Income Property, *Real Estate Appraiser and Analyst,* 1994, 50:1, 32–36.

Graham, M. F. and D. S. Bible, Classifications for Commercial Real Estate, *Appraisal Journal,* 1992, 60:2, 237–46.

Marks, A. P., C. Stanley and G. I. Thrall, Criteria and Definitions for the Evaluation of Geographic Information System Software for Real Estate Analysis, *Journal of Real Estate Literature,* July 1994, 2, 227–44.

Sirmans, G. S. and C. F. Sirmans, Property Manager Designations and Apartment Rent, *Journal of Real Estate Research,* 1991, 7:1, 91–98.

—— and J. D. Benjamin, Determining Apartment Rent: The Value of Amenities, Services and External Factors, *Journal of Real Estate Research,* Summer 1989, 4:2, 33–43.

Sullivan, A. M., *Urban Economics,* Homewood, Ill.: Irwin, 1990, 380.

Thrall, G. I. and A. P. Marks, Functional Requirements of a Geographic Information System for Performing Real Estate Research and Analysis, *Journal of Real Estate Literature,* January 1993, 1, 49–61.

Thrall, G. I., C. F. Sidman, S. Elshaw Thrall and T. J. Fik, The Cascade Diffusion Model for Measuring Housing Absorption by Small Area with a Case Study of St. Lucie County, Florida, *Journal of Real Estate Research,* 1993, 8:3, 401–20.

An earlier version of this paper was presented at the 1994 American Real Estate Society Meeting. The authors would like to acknowledge the helpful comments of the anonymous referees.