Assessing the Rental Value of Residential Properties: An Abductive Learning Networks Approach

*Kee S. Kim**
*Walt A. Nelson**

***Abstract*. This paper attempts to estimate rental value of residential properties using Abductive Learning Networks (ALN), an artificial intelligence technique. The results indicate that the ALN model provides an accurate estimation of rents with only seven input variables, while other multivariate statistical techniques do not. The ALN model automatically selects the best network structure, node types and coefficients, and therefore it simplifies the maintenance of the model. Once the final model is synthesized, the ALN model becomes very compact, rapidly executable and cost-effective.**

Introduction

Assessing the rental value of residential properties is a complex and challenging process to both practitioners and academicians because it involves analyzing the rental property, neighborhood characteristics and market conditions. The rental housing market is characterized as imperfect and inefficient, because the product is long-lasting, fixed on a given site, heterogeneous, and controlled by extensive governmental regulation. Since each rental housing market is confined to a given area, characteristics of a market in one area are not necessarily an accurate representation of other markets.

The purpose of this study is to build a model that can provide an accurate way of assessing the market value of residential rental property and analyzing the factors that determine market rents by using an artificial intelligence technique. Our main concern is to examine whether the Abductive Learning Networks (ALN) technique can be used to overcome many difficulties associated with the multiple regression technique which has been used primarily to analyze the price behavior of rental houses in the current literature (Guntermann, 1987; Murphy, 1989; Meacham, 1988; Jud and Winkler, 1991, Weirick and Ingram, 1990). The multiple regression technique is parametric and requires the user to specify the functional form of the solution. If one does not know or cannot guess the correct underlying form of the functional relationship, the regression approach will result in inaccurate models. Although the regression approach allows the use of a very general polynomial equation when the functional relationship is likely to be nonlinear, it becomes virtually impossible to estimate all coefficients, primarily because the number of coefficients to be estimated grows factorially as the number of variables and degrees of the function increase. Although the regression technique also allows the use of various transformations such as the Box-Cox (1964) transformation on the given data to adjust

*Southwest Missouri State University, 901 S. National Ave, Springfield, Missouri 65804.
Date Revised—September 1994; Accepted—July 1995.

for the nonlinearity of the dataset, the transformation can obscure the fundamental interconnections between variables depending on the nature of the dataset.

The survey data that are used for analyzing the rental price are often inaccurate, incomplete and distorted, and therefore, the outcome of the prediction might also be inaccurate within the framework of multivariate statistical analysis. An ALN may be a more appropriate technique, since it is nonparametric by nature and can learn and generalize the knowledge obtained from the relationship between the input and output, which are frequently changing over time, nonlinear, incomplete and/or unclear. Especially, neural networks on which the ALN is based ignore a relatively large amount of undesirable noise in input data set (Hinton, McClelland and Rumelhart, 1986), and can therefore provide better results.

In the following section, the Abductive Learning Network (ALN) is briefly discussed, and the theoretical framework for predicting rental prices is also presented. After describing the dataset, the model is constructed, and empirical results are presented and discussed in conjunction with results of the multiple regression model.

Abductive Learning Networks

The Abductive Learning Networks (ALN) technique (Barron, Mucciardi, Cook, Craig, and Barron, 1984) was developed from almost three decades of statistical modeling, neural network and artificial intelligence research. About forty years ago, scientists began to search for inductive algorithms to mimic what the central nervous system does best: associative reasoning, learning and thought. The ALN technique, evolved from the Group Method of Data Handling (GMDH) technique that was developed originally in Russia by Ivakhnenko (Hess and Montgomery, 1988), automatically generates the structure of the empirical model from the database and performs a traditional task of fitting model coefficients to bases of observational data. It uses mathematical functions in order to represent for numeric knowledge and it uses an artificial neural network structure to simplify the task of learning functional models by subdividing complex problems into smaller ones. The network consists of a number of processing units and interconnections between the units organized in a way that resembles neurons and synapses of a human brain. The power of the network lies in its ability to decompose complex problems into much smaller and simpler ones, and to solve them.

Montgomery (1989) developed an algorithm, called the Abductory Induction Mechanism (AIM™) which is a very effective computer-based algorithm for inductively creating abductive networks as models. AIM is a supervised inductive learning tool for automatically synthesizing models in the form of networks from a database of input and output variables. It was developed from the research of Barron (1984) and Montgomery (1989), and is based on "abducting reasoning," that is a "process of reasoning from a set of general principles to specifics *under uncertainty* using numeric functions and measures," and machine learning techniques called "Abductive Modeling™." The abductive model obtained from the AIM synthesis process is a layered network of feed-forward functional elements, in which the coefficients, number and types of network elements, and the connectivity are learned inductively and automatically. The abductive network consists of arcs and nodes, and each node has unique multivariable configurations: singles, doubles, triples, normalizers, white elements, unitizers, and wire elements (Montgomery, 1989).

$$Single = W_0 + W_1*X_1 + W_2*X_1^2 + W_3*X_1^3 ,$$

$$Double = W_0 + W_1*X_1 + W_2*X_2 + W_3*X_1^2 + W_4*X_2^2 + W_5*X_1*X_2 + W_6*X_1^3 + W_7*X_2^3,$$

$$\begin{aligned} Triple = {} & W_0 + W_1*X_1 + W_2*X_2 + W_3*X_3 + W_4*X_1^2 + W_5*X_2^2 + W_6*X_3^2 \\ & + W_7*X_1*X_2 + W_8*X_1*X_3 + W_9*X_2*X_3 + W_{10}*X_1*X_2*X_3 \\ & + W_{11}*X_1^3 + W_{12}*X_2^3 + W_{13}*X_3^3 , \end{aligned}$$

where X_i and W_i denote input variables and coefficients respectively. These elements are homogeneous of multinomial of degree 3 in one, two, three variables and allow interaction among input variables. Each node in a network is represented by one of the above equations, in which only significantly contributing terms appear.

Normalizers use the function, $W_0 + W_1X_1$, to transform the original input variables into standardized normal variables with a mean of zero and a variance of one. The white element is $W_1X_1 + W_2W_2 + \ldots + W_nX_n$, which is a linear combination of all inputs to the current layer. Unitizers convert the normalized data back into the original data to assess the output data, while the wire element is used for a network that consists only of a normalizer and a unitizer. The output of elements in one layer can feed subsequent layers, together with the original input variables to synthesize networks from layer to layer until the network model ceases to improve.

The objective of the ALN algorithm is to train and identify a model that minimizes the predicted squared error (*PSE*), the errors on as yet unforeseen data, without overfitting the data (A. R. Barron, 1984). *PSE* consists of the training squared error (*TSE*) and overfit penalty, $2\sigma^2K/N$, as shown in the following:

$$PSE = TSE + 2\sigma^2 \frac{K}{N}, \tag{1}$$

where *TSE* is the average squared error of the model on the training sample observations, K is the number coefficients that are estimated to minimize *TSE*, σ_p^2 is a prior estimate of true error variance, and N is the size of the training sample observations. The training squared error (*TSE*) decreases by nature at a decreasing rate, but the overfit penalty increases linearly, as each coefficient (K) is added to the model. If the abductive model is obtained by minimizing *TSE* alone, the model will likely perform well on the training dataset, but it will perform poorly on future observations or hold-out samples, especially when the model has an overly complex structure that requires to estimate many coefficients. By adding a term for overfit penalty, we can be certain to obtain the minimum expected squared difference between the estimated model and the true model on future dataset (A. R. Barron, 1984).

The Model and Data Description

Rent prediction is usually modelled as a function of a bundle of housing characteristics, called supply-side attributes, which describe the unit supplied by the landlord (Sirmans and Benjamin, 1991). These characteristics typically include clusters of variables such as dwelling size, dwelling amenities and dwelling location. The numbers of rooms and bathrooms (which describe dwelling size) are typically associated with increase in rent (Barnett, 1979; Engle and Marshall, 1983; Follain and Malpezzi, 1980; Noland, 1980; Shear, 1983), but the number of bedrooms is usually very insignificant in the regression model.

Other supply-side attributes include amenities and building characteristics. Amenities such as appliances (Noland, 1980; Follain and Malpezzi, 1980) and central air conditioning (Follain and Malpezzi, 1980) are usually associated with higher rent levels. Although building size (number of units) is usually associated with higher rent levels (Follain and Malpezzi, 1980; Barnett, 1979), building age is usually associated with lower rent levels (Noland, 1980; Engle and Marshall, 1983; Follain and Malpezzi, 1980). Additionally, Porell (1985) and Noland (1980) show that the presence of on-site landlords or managers is associated with lower rent levels.

Although the presence of abandoned buildings and adequate shopping facilities and schools, is sometimes considered a supply-side attribute, these variables are often treated separately from dwelling attributes in rent models and considered as neighborhood characteristics. Shear (1983) and Follain and Malpezzi (1980) show that abandoned buildings, lack of adequate schools and shopping facilities are associated with lower rent levels.

Frequently, the rent model is augmented by a group of demand-side characteristics describing the tenant and his household. Muth (1985) suggests that tenant preferences and demographics be taken into consideration when developing a rent model. Shear (1983) shows that tenant age, income and education level all contribute to higher levels of occupancy turnover, meaning such households are more mobile. Follain and Malpezzi (1980) show that minority-group status is associated with lower levels of rent, perhaps due to lack of mobility for such households.

The model in this paper incorporates all variables found to be significant in past research, as well as other theoretically sound variables. These independent variables are organized into four categories as follows:

$$RENT = f(BC, LC, TC, NC)\,, \tag{2}$$

where

BC = Building Characteristics;
LC = Landlord Characteristics;
TC = Tenant Characteristics;
NC = Neighborhood Characteristics.

The entire set of input variables is explained in Exhibit 1 and their definitions are presented in Appendix A. Building characteristics are subdivided into building type, size of unit, amenities, and maintenance. Building type includes the number of units in the structure and age of the building. Size of rental unit includes number of rooms, number of baths and number of bedrooms. Amenities include plumbing facilities, kitchen appliances, utilities paid, and central air conditioning. Maintenance includes the presence of exterior leaks and broken windows. Landlord characteristics include the investor's age, years of experience and number of properties owned. However, the database used in this study includes one landlord variable—*LLBLG*—that indicates whether the landlord or his manager lives on the premises.

Exhibit 1
Variable Definition

Building Characteristics			
	Size of Unit	*BEDRMS*	Number of bedrooms
		BATHS	Number of bathrooms
		ROOMS	Number of rooms
	Amenities	*APPIND*	Number and age of appliances
		HEATING	Heat paid by landlord*
		CNTRLAC	Central air conditioning*
	Maintenance	*LEAK*	Exterior leaks*
		PLUMB	All plumbing facilities present*
		EBROKE	Broken windows*
	Structure	*NUNITS*	Number of apartment units
		BLDAGE	Age of property
Landlord Characteristics			
		LLBLG	On-site landlord or manager*
Tenant Characteristics			
		LOR	Tenant length-of-residence
		CROWDS	Ratio of persons in household to number of rooms
		HEADAGE	Age of head-of-household
		*GRADE*1	Education level of head-of-household
		ZINC	Household income
		BRACE	Head-of-household is Black*
		*SPAN*1	Head-of-household is Hispanic*
Neighborhood Characteristics			
		ABNDON	Abandoned buildings nearby*
		CRIME	Neighborhood crime bothers tenant*
		LITTER	Neighborhood litter bothers tenant*

*indicates a categorical (0,1) variable

Tenant characteristics include the head-of-household age, race and education level. Also included are household income, number of children and length-of-residence. Neighborhood characteristics include the tenant's opinion of the neighborhood and whether or not crime is perceived to be a problem. Additional variables included are the presence of abandoned buildings and noticeable litter.

Both the multiple regression and ALN models use twenty-two variables as input variables and one output variable for training as shown in Exhibit 1. The total sample is randomly divided into one for training and another for evaluating the model. The ALN model is trained by using the AIM™ package that is developed and distributed by the AbTech Corporation.

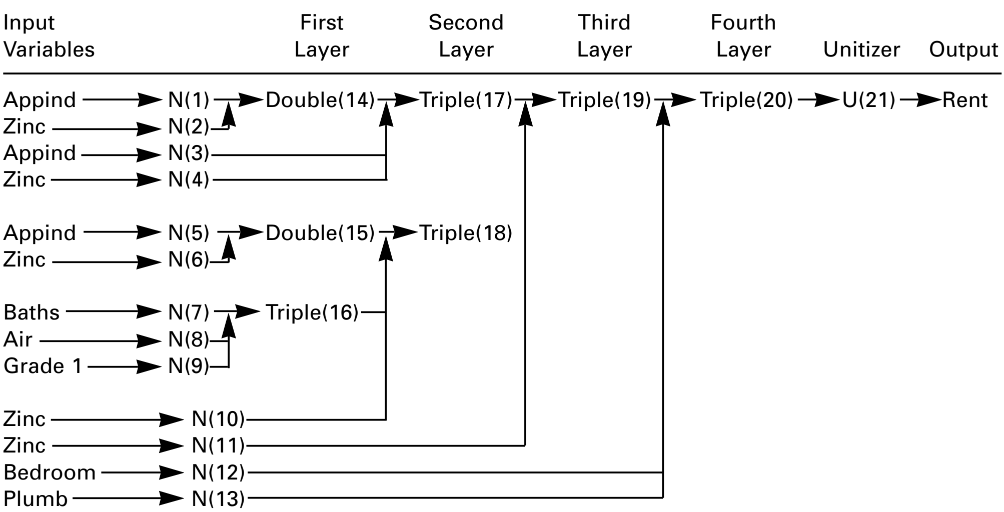
The Atlanta sample includes 1,031 rental dwelling units retrieved from the 1987 American Housing Survey (AHS) sponsored by the Department of Housing. The sample contains moderately priced two-bedroom apartments the rent for which is no more than $700 per month. Most of the units are located in older buildings having five or fewer units. There is usually no on-site landlord or manager. The typical head-of-household is a high school graduate, aged 40, and White. The survey also shows that most tenants have lived in their residence one-to-three years, and hold a moderately high opinion of the quality of their neighborhood.

Empirical Results

An Abductive Learning Network (ALN) is synthesized in Exhibit 2 from the training dataset, using the AIM™ software that is developed by the AbTech Corporation. It is a robust and compact transformation structured as a layered network of feedforward functional elements. It contains the best network structure, node types, coefficients, and connectivity to minimize the predicted squared error (PSE) without outfitting the data. It includes seven different input variables that contribute significantly to rents determination, and some of the selected input variables are used repeatedly in the network. ALN also includes nodes such as Nomalizers, Doubles, Triples, and Unitizers as a part of the network, and each node is represented by an equation with estimated coefficients.

The equations in Exhibit 3 represent individual nodes in the network shown in Exhibit 2; and each equation number corresponds to the node number in the network (ALN). In this network seven input variables are first transformed into standardized normal variables with the mean of zero and a variance of one using normalizers in Exhibit 3; an *APPIND* variable, for example, is normalized by equation 1. Once the seven input variables are normalized, they are fed into the first layer to generate a series of intermediate output values. For example, equation 14 uses normalized values of *APPIND* and *ZINC* variables as input values to generate an intermediate output. This intermediate output from *DOUBLE* (14) and two other inputs are fed again into the node, *TRIPLE*(17), as an input variable in order to generate an output in the second

Exhibit 2
Abductive Network Model of the ALN Approach



Note: The numbers in the parentheses correspond to the equation numbers as shown in Exhibit 3.

Exhibit 3
Abductive Network Equations

Normalizers

1. *APPIND* = $-.502+1.54X_1$
2. *ZINC* = $-.641+.659X_1$
3. *APPIND* = $-.502+1.54X_1$
4. *ZINC* = $-6.41+.659X_1$
5. *APPIND* = $-.502+1.54X_1$
6. *ZINC* = $-6.41+.659X_1$
7. *BATH* = $-2.31+1.75X_1$
8. *AIR* = $-1.55+2.2X_1$
9. *GRADE*1 = $-4.14+3.19X_1$
10. *ZINC* = $-6.41+.659X_1$
11. *ZINC* = $-6.41+.659X_1$
12. *BDRM* = $-2.31+1.15X_1$
13. *PLUMB* = $-27.1+27.1X_1$

Doubles

14. *DOUBLE* = $.193+4.92X_1+3.42X_2+.152X_1^2+.37X_2^2 +2.89X_1X_2-.000187X_1^3+.0031X_2^3$
15. *DOUBLE* = $.193+4.92X_1+3.42X_2+.152X_1^2+.37X_2^2 +2.89X_1X_2-.000187X_1^3+.0031X_2^3$

Triples

16. *TRIPLE* = $.18+.589X_1+.0855X_2+.373X_3-.104X_1^2 -.263X_3^2+.0771X_1X_3-.0785X_2X_3-.139X_1X_2X_3 +.0044X_1^3+.0914X_2^3-.0154X_3^3$
17. *TRIPLE* = $22.7-115X_1+574X_2+400X_3-.443X_1^2 +17.5X_2^2+40.8X_3^2+1.95X_1X_2 +2.81X_1X_3+336X_2X_3+.739X_1X_2X_3 +.0195X_1^3-.0173X_2^3$
18. *TRIPLE* $=.0766+1.3X_1+.203X_2-.44X_3-.0749X_1^2 +.0424X_2^2-.61X_3^2+.0892X_1X_2+.659X_1X_3 -.127X_2X_3+.0459X_1X_2X_3-.0153X_1^3 -.0671X_2^3-.0788X_3^3$
19. *TRIPLE* = $-.26+1.9X_1+.233X_2-.0568X_3+.0704X_1^2 +.129X_2^2-.0291X_3^2-.174X_1X_2+.198X_1X_3 -.154X_2X_3+.0779X_1X_2X_3-.0316X_1^3 -.0564X_2^3-.00283X_3^3$
20. *TRIPLE* = $.998X_1+.0144X_2+.000882X_3^2$

Unitizers

21. *RENT* = $40.6+140X_1$

layer. In the same fashion, these intermediate output values in the second layer are fed into the subsequent layers to estimate the final output value in the fourth layer from the node *TRIPLE*(20). This value is converted back to rental values with the mean and variance of the original output values that are used to train the network, using a unitizer, *U*(21). This network now becomes a knowledge base from which a series of rental values can be estimated from the seven input variables.

The results of rent estimation by the ALN and regression approaches are shown in Exhibit 4. Although the table shows only the first 38 out of 245 cases we tested, it is an

Exhibit 4
Rent Estimation by the ALN and REG Approaches*

Obs	Rent	ALN Pred	REG Pred	ALN Err	REG Err
1	108	104.36	277.52	−3.64	169.52
2	111	231.2	383.15	120.20	272.15
3	250	255.15	501.39	5.15	251.39
4	270	265.9	239.11	−4.10	−30.89
5	270	264.43	313.92	−5.57	43.92
6	300	298.46	397.13	−1.54	97.13
7	300	301.02	427.71	1.02	127.71
8	300	296.89	402.56	−3.11	102.56
9	305	301.41	371.49	−3.59	66.49
10	325	323.29	377.3	−1.71	52.30
11	325	328.58	465.5	3.58	140.50
12	342	341.94	467.3	−.06	125.30
13	345	339.96	368.89	−5.04	23.89
14	350	346.1	440.01	−3.90	90.01
15	350	347.66	482.87	−2.34	132.87
16	350	345.68	451.09	−4.32	101.09
17	350	346.77	390.31	−3.23	40.31
18	360	363.09	382.26	3.09	22.26
19	360	365.17	411.49	5.17	51.49
20	365	365.47	463.22	.47	98.22
21	371	421.66	451.55	50.66	80.55
22	375	368.35	391.69	−6.65	16.69
23	375	368.97	365.71	−6.03	−9.29
24	380	382.17	419.67	2.17	39.67
25	380	376.3	410.82	−3.70	30.82
26	380	372.78	418.79	−7.22	38.79
27	380	375	434.99	−5.00	54.99
28	385	378.42	395.93	−6.58	10.93
29	385	391.49	382.99	6.49	−2.01
30	390	390.89	464.91	.89	74.91
31	395	395.5	422.92	.50	27.92
32	395	402.11	401.38	7.11	6.38
33	400	401.97	564.15	1.97	164.15
34	405	400.71	385.24	−4.29	−19.76
35	415	412.67	498.4	−2.33	83.40
36	415	421.24	425.88	6.24	10.88
37	420	414.36	452.23	−5.64	32.23
38	425	430	403.93	5.00	−21.07

*The prediction of the first 38 observations out of 245 is shown for convenience.

adequate representation of total sample. The result indicates that the ALN model outperforms the regression such that the ALN model generates very accurate and consistent estimation of rents while the regression model is not able to do so. The scatter diagram in Exhibit 5 shows graphically how accurate the ALN approach is relative to the regression approach for estimating market value of rental houses.

The descriptive statistics of rent estimation on the evaluation sample in Exhibit 6 also show that the average absolute error ($7.92) of estimation by the ALN is much less than

Exhibit 5
An Assessment of Rents by the ALN and Regression Approaches

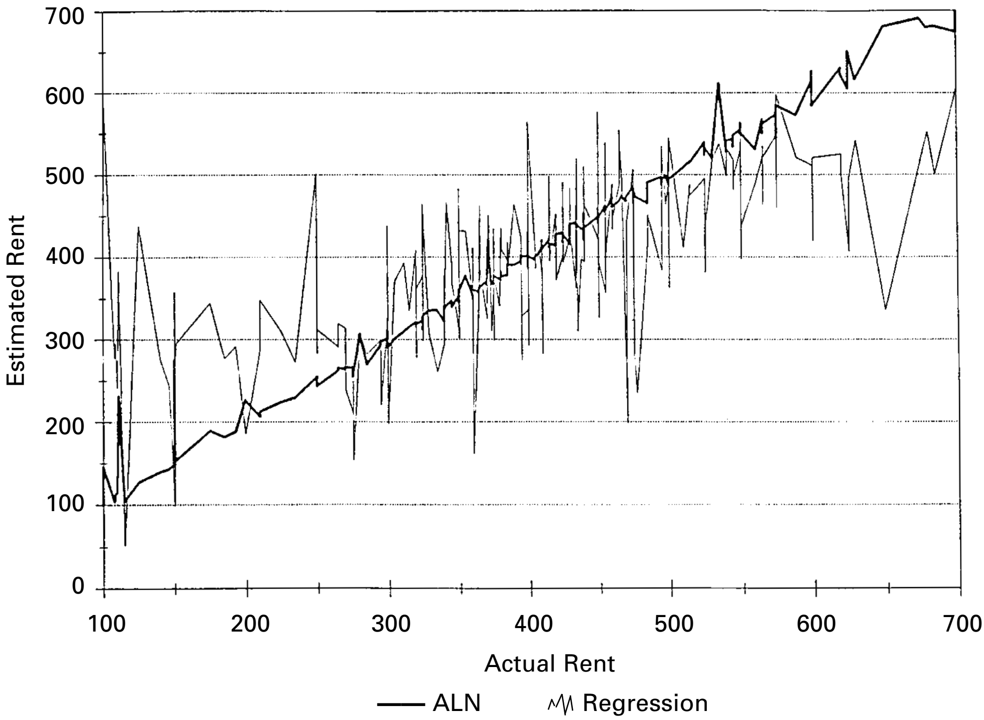


Exhibit 6
Descriptive Statistics of Valuation on the Evaluation Sample

	Market Value	ALN Value	ALN Error	REG Value	REG Error
Sample	245	245	245	245	245
P-Corr	1.00	.991	N/A	.692	N/A
Mean	403.22	404.67	7.92*	398.87	71.03*
Std Dev.	133.25	131.53	13.53*	98.964	65.01*
Min	100.00	100.00	−30.46	52.833	−498.19
Max	700.00	700.00	125.61	605.37	314.06
Mad	77.00	81.31	3.97	74.26	54.43

the average error ($7.103) by the regression approach. In addition, the Pearson correlation between the actual and estimated rents by the ALN is very high (.991) relative to that (.692) of the regression model. This indicates that the rents estimated by the ALN have not only a lower average deviation but also move very closely together with the actual rent.

The critical variables that contribute significantly to the determination of the rental value are shown in Exhibit 7. The ALN model selected seven out of twenty-two input variables as significant variables while the regression model (Appendix B) selected ten input variables. Exhibit 8 also indicates that the rent estimated by the ALN model is very sensitive to changes in the appliance (*APPIND*) and household income (*ZINC*) variables. For example, at the given input level (second column), the estimated rent is $459.51 according to the ALN model. If the value of the *ZINC* increases 10% to 10.689, the estimated rent increases 52.34% over $459.51 to $700. When the *ZINC* value decreases by 10%, the rent also decreases by 64.14% to $164.82 with all other variables being constant. The fact that the rental value is very sensitive to the appliance variable (*APPIND*) implies that tenants may be willing to pay a premium for newer and

Exhibit 7
Value Determinants of the AIM and Regression Models

	Variable Names	ALN Approach Critical Variables	REG Approach Critical Variables
1	*BEDRMS*	YES	
2	*BATHS*	YES	YES
3	*LEAK*		
4	*ROOMS*		YES
5	*PLUMB*	YES	
6	*EBROKE*		
7	*HEATINC*		
8	*CNTRLAC*	YES	YES
9	*NUNITS*		
10	*APPIND*	YES	YES
11	*BLDAGE*		YES
12	*LLBLG*		
13	*LOR*		YES
14	*CROWDS*		
15	*HEADAGE*		YES
16	*GRADE*1	YES	YES
17	*ZINC*	YES	YES
18	*BRACE*		YES
19	*SPAN*1		
20	*ABNDON*		
21	*CRIME*		
22	*LITTER*		

Exhibit 8
Sensitivity of Rents to the Input Change

		10% Increase		10% Decrease	
Estimated Variable	Current Input	Estimated Rents	% Change	Estimated Rents	% Change
BEDRMS	1.997	459.98	.10	459.04	−.10
BATHS	1.320	459.63	.03	459.40	−.02
PLUMB	1.001	460.42	.20	459.51*	.00
CNTRLAC	.708	459.52	.00	459.52	.00
APPIND	.325	507.19	10.38	411.25	−10.50
*GRADE*1	12.973	459.65	.03	459.39	−.03
ZINC	9.717	700	52.34	164.82	−64.13
Rent	459.51				

*Rent was calculated at 1.0 the *PLUMB* variable.

a greater number of appliances. The results of the model also indicate that among other variables the household income (*ZINC*) becomes a very important factor for determining rents.

Once the ALN model is built, it can be used to assess rents by using seven input variables. The ALN model now becomes an expert that is very dynamic and robust for estimating rents, because the ALN model responds differently to the change in an input variable, depending upon not only the values of the variable involved but also of other input variables. The results of the ALN model can now be effectively utilized for practitioners to make market-derived rent adjustments, to set rents and to design apartments projects.

Concluding Remarks

This paper discusses the successful building of an ALN model from survey data that was capable of generating an accurate estimate of rental values. Although the dataset that consists of system-level inputs and an output included normal as well as abnormal (outlier) data, the ALN model was able to induce from the data the knowledge necessary to estimate rental values accurately and automatically. This research also demonstrated that the ALN model is superior to the regression model that has commonly been used in the area of estimating rents.

The ALN technique is a powerful supervised inductive learning tool for synthesizing an abductive network model that can reveal a subtle relationship that is not otherwise apparent. Since the AIM™ package automatically selects the best network structure, node types and coefficients, it simplifies the maintenance of the model such that a new model can be synthesized to incorporate an additional or new dataset when new knowledge or information is warranted. Above all, the ALN model created by the AIM package is very compact, rapidly executable and cost-effective.

Appendix A
Variable Definitions

BEDRMS	*Number of bedrooms*		
	0	=	efficiency unit
	1–9	=	1 to 9 bedrooms
	10	=	10 or more bedrooms
BATHS	Number of bathrooms		
LEAK	Tenant noticed an exterior leak in the past 12 months		
	1	=	Yes
	2	=	No
ROOMS	Number of rooms in unit		
	1–20	=	1 to 20 rooms
	21	=	21 or more rooms
PLUMB	Unit has complete plumbing facilities		
	1	=	Yes
	2	=	No, lacks a few items (i.e. sink, tub)
	3	=	No, lacks all items
EBROKE	Unit has broken windows		
	1	=	No
	2	=	Yes
HEATING	Heat is included in rent		
	0	=	No
	1	=	Yes
CNTRLAC	Unit has central air conditioning		
	0	=	No
	1	=	Yes
APPIND	The appliance index is the sum of two components. One component represents the presence of a particular appliance, while the other is an estimate of depreciation for older appliances, if present. The value of 1.0 is added to the index if any of the following appliances are present: disposal, refrigerator, stove, dishwasher, clothes washer, or clothes dryer. The value of .5 is subtracted from the index if any of the appliances listed above is at least five years old. Our sample does not indicate dishwashers older than 5 years.		
NUNITS	Number of units in structure		
	1–1000	=	1–1000 units
	1001	=	1001 or more units
BLDAGE	Age of the rental property (in years)		
	1	=	built in 1987
	2–9	=	built between 1979–86
	12	=	built 1975–78
	16	=	built 1970–74
	23	=	built 1960–69
	33	=	built 1950–59
	43	=	built 1940–49
	53	=	built 1930–39
	63	=	built 1920–29
	73	=	built before 1920

LLBLG	Landlord or manager lives on premises 1 = Yes 2 = No
LOR	Tenant length-of-residency in years
CROWDS	Ratio of persons in household to rooms in unit. Illus.: 1 person to 5 rooms = .20 or "roomy" 2 persons to 5 rooms = 2.0, or "crowded"
HEADAGE	Age of head-of-household
*GRADE*1	Education level of head-of-household 1–12 = 1st grade through high school graduate 13–16 = college education 17–18 = graduate education
ZINC	Annual household income
BRACE	Head-of-household is 0 = White, Oriental, or Hispanic-White 1 = Black or Hispanic-Black
*SPAN*1	Head-of-household is 0 = Not Hispanic 1 = Hispanic
ABNDON	Surveyor observed abandoned buildings nearby 0 = No, none observed 1 = Yes
CRIME	Tenant indicates neighborhood crime is bothersome 0 = No, not bothersome 1 = Yes
LITTER	Tenant indicates litter in neighborhood is bothersome 0 = No, not bothersome 1 = Yes

Appendix B
Unweighted Least Squares Linear Regression of Rent

Predictor Variables	Coefficient	Std Error	Student's *t*	*P*-Value
Constant	−186.043	117.042	−1.59	.1124
BEDRMS	13.2195	7.62173	1.73	.0833
*BATHS**	53.0867	7.24319	7.33	.0000
LEAK	−.50072	5.44248	−.09	.9267
*ROOMS**	16.4360	4.50341	3.65	.0003
PLUMB	145.032	102.786	1.41	.1587
EBROKE	25.9022	45.4296	.57	.5687
HEATING	29.2059	16.8850	1.73	.0841
*CNTRLAC**	74.0209	10.9246	6.78	.0000
NUNITS	−.56770	.68641	−.83	.4085
*APPIND**	14.6455	5.92615	2.47	.0137
*BLDAGE**	−.83800	.27973	−3.00	.0028
LLBLG	7.38428	8.67527	.85	.3950
*LOR**	−4.44809	1.02175	−4.35	.0000
CROWDS	−26.0980	16.1833	−1.61	.1073
*HEADAGE**	.67432	.28761	2.34	.0193
*GRADE*1*	11.2955	1.37905	8.19	.0000
*ZINC**	9.61141	2.75729	3.49	.0005
*BRACE**	−9.88610	4.78860	−2.06	.0393
*SPAN*1	15.2369	22.4515	.68	.4976
ABNDON	6.84152	20.5142	.33	.7389
CRIME	3.79072	15.3462	.25	.8050
LITTER	−27.4741	19.2171	−1.43	.1533

R-sq.	.5289	Resid. Mean Sq. (MSE)	9555.37
Adj. *R*-sq.	.5143	Std Dev.	97.7516

Source	DF	SS	MS	*F*-Value	*P*-Value
Regression	22	7.607E+06	3.458E+05	36.19	.0000
Residual	709	6.775E+06	9555.37		
Total	731	1.438E+07			

Cases included	732	Missing cases	0

*independent variables significant at 5% level

References

AIM User's Manual, AbTech Corporation, Charlottesville, Va.: 1993.

Barnett, L., *Using Hedonic Indexes to Measure Housing Quality,* Santa Monica, Calif.: Rand Corporation, R-2450-HUD, 1979.

Barron, A. R., Predicted Squared Error: A Criterion for Automatic Model Selection, in S. J. Farlow, editor, *Self Organizing Methods in Modeling: GMDH Type Algorithms,* New York: Marcel Dekker, 1984.

Barron, R. L., et al., Adaptive Learning Networks: Development and Application in the U.S. of Algorithms Related to GMDH, in S. J. Farlow, editor, *Self-Organizing Methods in Modeling: GMDH Type Algorithms,* New York: Marcel Dekker, 1984.

Box, G. E. P. and D. R. Cox, An Analysis of Transformations, *Journal of the Royal Statistical Society,* 1964, B26, 211–43.

Engle, R. F. and R. C. Marshall, A Microeconometric Analysis of Vacant Housing Units, in J. C. Weicher, K. E. Villani and E. A. Roistacher, editors, *Rental Housing: Is There a Crisis?,* 105–23, Washington, D.C.: Urban Institute Press, 1983.

Follain, J. R., Jr. and S. Malpezzi, *Dissecting Housing Value and Rent,* Washington, D.C.: Urban Institute, UI 249-17, 1980.

Guntermann, K. L. and S. Norrbin, Explaining the Variability of Apartment Rents, *AREUEA Journal,* 1987, 15:4, 321–40.

Hess, P. and G. J. Montgomery, Abduction: Theory and Application, *Proceedings of the 4th Annual Aerospace Applications of Artificial Intelligence Conference (AAAIC),* October 1988.

Hinton, G. E., J. L. McClelland and D. E. Rumelhart, Distributed Representation, in *Parallel Distributed Processing: Exploration in the Microstructure of Cognition,* 77–109, Cambridge, Mass.: MIT Press, 1986.

Jud, G. D. and D. T. Winkler, Location and Amenities in Determining Apartment Rents: An Integer Programming Approach, *Appraisal Journal,* April 1991, 266–75.

Lippmann, R., An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine,* April 1987, 4–22.

Meacham, A., Applying Regression Analysis to Real Estate Appraisals, *Real Estate Appraiser and Analyst,* Summer 1988, 23–27.

Montgomery, G. J., Abductive Diagnostics, *Proceedings of the AIAA Computers in Aerospace Conference,* Monterey, California, October 1989.

Murphy, L., T., III, Determining the Appropriate Equation in Multiple Regression Analysis, *Appraisal Journal,* October 1989, 498–517.

Muth, R. F., Models of Land-Use, Housing, and Rent: An Evaluation, *Journal of Regional Science,* November 1985, 593–606.

Noland, C. W., *Assessing Hedonic Indexes for Housing.* Santa Monica, Calif.: Rand Corporation, N-1505-HUD, 1980.

Porell, F. W., One Man's Ceiling is Another Man's Floor: Landlord/Manager Residency and Housing Condition, *Land Economics,* 1985, 61, 106–117.

Quang, D. A. and G. Grudnitski, A Neural Network Approach to Residential Property Appraisal, *Real Estate Appraiser,* December 1992, 38–45.

Shear, W. B., A Note on Occupancy Turnover in Rental Housing Units, *AREUEA Journal,* 1983, 11, 525–38.

Sirmans, G. S. and J. D. Benjamin, Determinants of Market Rent, *Journal of Real Estate Research,* 1991, 6:3, 357–78.

Smith, L. B., K. T. Rosen and G. Fallis, Recent Developments in Economic Models of Housing Markets, *Journal of Economic Literature,* 1988, 26, 29–64.

Weirick, W. N. and F. J. Ingram, Functional Form Choice in Applied Real Estate Analysis, *Appraisal Journal,* January 1990, 57–73.