

SUBMARKETS IN URBAN HOUSING

BY

→ ARDEN BRUMMELL

AND

DAVID KARP

MCMASTER UNIVERSITY
HAMILTON, ONTARIO

FEBRUARY, 1983

REPORT PREPARED FOR
CANADA MORTGAGE AND HOUSING CORPORATION
CONTRACT NUMBER 6585/B9

ACKNOWLEDGEMENTS

We would like to acknowledge the work of Michelle Elliot and Barb Konyi in the onerous task of compiling and coding the data used in this study and the helpful suggestions of Phillip Brown of CMHC in developing the study and improving the result.

We would also like to acknowledge the work of Donna MacLeod in cheerfully typing and revising this manuscript over a lengthy period of time.

SUBMARKETS IN URBAN HOUSING

I. INTRODUCTION

Market Factors and Property Factors

In understanding urban housing markets, an important problem is to understand the factors influencing house prices. Two broad factors may be identified: market factors and property factors. Market factors refer to factors which influence the supply and demand for housing. These include influences such as changes in income, rates of household formation, interest rates, etc. The problem is to understand differences in prices over time or between markets. Some aggregate measure of price such as the average or medium value is often used as the prime indicator.

Property factors refer to factors which influence the prices of individual properties. Typically, the basic hypothesis is that the price of a residence depends on the set of attributes present or associated with that residence, e.g., floor space, age, lot size, type of neighbourhood, etc. The problem is to understand price differences between properties within a given market. Property factors, therefore, focus on price variations within a market while market factors focus on differences between markets (either in space or time). The two sets of factors, however, are not independent.

In analyses of either type it is often implicitly assumed that the market is well defined. In studies analyzing differences between distinct cities there appears to be little conceptual difficulty. But even in relatively simple comparisons the definition of the market may be difficult. For example, is Hull part of the Ottawa housing market or not? Similarly, to what extent can the housing market in Quebec City, an old city with a corresponding array of older properties, be compared to the housing market in Calgary which on average is much newer. These differences, of course, are recognized and reflected in the comparisons conducted by national real estate firms which compare prices for specific types of housing in specific types of neighbourhoods. In essence, this approach attempts to remove the effects of property factors in comparing prices between markets.

Nature of Submarkets

A similar approach is necessary in analyzing price differences within an urban housing market. MacLennan (1977) has argued that the housing market in an urban area is not a single well connected market but a series of quasi-independent submarkets. Submarkets are thought to arise because both the supply and demand for housing appear to be highly segmented. Housing is a complex, heterogeneous commodity. As a result, the supply of housing is not simply the total number of vacancies but the number of vacancies of each kind at a given moment in time. Similarly the demand for housing may be segmented. Households in different life cycle stages, with different incomes have different needs and preferences and generate different

demands. Submarkets occur, therefore, if different types of households tend to compete for different types of housing with little substitution between housing types.

The impact of this type of structure is to create a disjointed market which responds to change in a complex way. Consider as an analogy the movements of a rope and chain. A rope is continuous. Movement of the rope at any one point affects all other parts in a smooth and continuous fashion. This represents a well connected market. A segmented market, in contrast, operates like a chain. Movement of one link in the chain may or may not affect adjacent links depending on the extent of the movement. With sufficient movement, however, neighbouring links will be affected sending out a jerky wave along the chain. In this way, each link, or submarket, has a certain degree of freedom to move independently of other links in the chain.

To put this in a housing context, consider the effect of an increase in demand for one segment of the market. For example, a rapid increase in the 25 to 35 age group due to migration might increase the demand for medium priced suburban bungalows. This would lead to an increase in prices for this type of housing. Existing households would be encouraged to move, probably to larger more expensive houses thereby forcing up the prices of those houses. Similarly households in less expensive houses would be blocked from moving up the housing ladder decreasing the supply and raising prices in that segment of the market. Depending on the supply of new

houses the price level of the entire market would be affected. This process is common whether the market is well connected or segmented. What differs is the timing and continuity of the transition. In a well connected market the time lags are small and prices move smoothly. In a segmented market, however, long time lags may develop and discontinuities appear. Existing households will be unwilling to substitute one type of housing for another so that the price of suburban bungalows may increase significantly and remain "out of step" for some time. Of course, some substitution will occur and eventually one link will exert an influence on adjacent links. The important result, however, is that at any moment in time, house price differences may occur both because of differences in the level of attributes and because of differences in submarket conditions.

Implications for Policy and Analysis

The implications of this type of market structure are important from two perspectives, policy and analysis. From a policy perspective, the implication of submarkets is that the effects of changes in overall market conditions, e.g., changing interest rates, may differ between submarkets. In the extreme one segment of the market could be experiencing increases in supply and falling prices while another is experiencing increases in demand and rising prices. This means that broad policies implemented in an urban area may have quite different impacts in different submarkets. Particularly, the magnitude and time lag of policy impacts are likely to

vary. The result is that formulating and evaluating housing policy in this type of volatile market is difficult without a firm understanding of the extent of market segmentation in urban housing.

The submarket hypothesis is also important from an analytical perspective. The hypothesis implies that the prices of individual houses will vary not only because the level of attributes vary across houses, but also because market conditions vary across submarkets. This means that the relationship between house prices and housing attributes may vary across submarkets, i.e. the relative importance of attributes may vary. This affects the method of estimating and analyzing the price attribute relationship. If a single total market is assumed, then an appropriate approach is to estimate the relationship using a sample pooled across house types, neighbourhood types and locational types. If this aggregate approach is used, however, and submarkets exist, then the estimated coefficients of the relationship, which reflect the implicit prices of the attributes (Rosen, 1974), will measure some average set of prices but not the accurate price in any specific submarket. If submarkets exist, then variations between submarkets must be accounted for. The simplest approach is to estimate the price-attribute relationship in each submarket separately. This, of course, greatly complicates the analysis. It is important, however, from an analytical perspective to examine the extent of market segmentation in urban housing and its impact on the price attribute relationship.

Purpose of Study

The purpose of the present study is to examine the hypothesis that urban housing markets are segmented into a number of submarkets. Essentially there are two questions: 1) do submarkets exist and 2) what is the magnitude and importance of their impact? In both cases the analysis depends on the implication of the submarket hypothesis that the presence of submarkets will lead to differences in the price attribute relationship. If such differences can be demonstrated, then this provides evidence of the existence of submarkets. The magnitude and importance of submarkets can then be evaluated by comparing the explanatory power, both statistically and logically, of the submarket and single market models. The logic of the analysis involves 4 steps: 1) define a set of potential submarkets, 2) estimate the price-attribute relationship within each submarket, 3) compare the estimates across submarkets, and 4) compare the estimated submarket and single market models.

Specific Aspects of the Study

While the logic of the analysis is straightforward, operationalizing the analysis involves a number of difficulties. At least two important problems may be noted. First, a well developed conceptualization of the key variables involved in housing market segmentation has not emerged as a basis for defining submarkets. Second, a well specified model of the price-attribute relationship has not been developed either with respect to the set of specific variables which should be included or with respect to

the form of the relationship. The result, therefore is that the analysis, like previous analyses, explores a number of submarket definitions and model specifications.

The present study differs from previous studies primarily in the range of variables used and the number of submarket divisions explored. Like many previous studies a large number of structural variables, i.e., specific dwelling characteristics are included. Unlike most previous studies, however, a number of both physical and social neighbourhood variables are included. The inclusion of social characteristics represents an effort to identify the effects of demand segmentation in generating submarkets. That is, do specific types of people tend to purchase similar types of housing? The large number of physical neighbourhood variables, particularly several land use measures, represents an effort to identify the effects of competing land uses on housing prices, a problem identified in an earlier study (Brummell, 1981). Finally, a somewhat different locational variable, distance to the major regional employment centre, is examined to capture the multi-locational nature of accessibility and the dominance of the industrial centre in Hamilton, Ontario, the city used in the analysis.¹

The present study also differs in the number of submarket divisions explored. The analysis explores potential submarkets on the basis of 1) house size (i.e. number of room), 2) neighbourhood housing age, 3)

¹ The accessibility variables are removed later in the analysis.

neighbourhood income, 4) spatial areas and 5) house price. Few previous studies, and no Canadian studies, have explored the concept of submarkets as extensively along these lines.

In the following section (II) the results of previous studies are reviewed. The submarket model is then presented (III) as the basis for the empirical analysis. The data and definition of the variables are then discussed (IV) followed by the empirical analysis (V). First the results for the total market are presented under the hypothesis of a single total market. Then the results for the specific submarkets are analyzed. Finally, a summary discussion and overall conclusions are presented (VI).

II PREVIOUS STUDIES

Previous studies of the price-attribute relationship may be divided into two groups. Both have been concerned with identifying the determinants of house prices, but one group has implicitly assumed a single market (e.g., Ball, 1973; Grether and Mieszkowski, 1974) while the other group has either been directly concerned with market segmentation (Schnare and Struyck, 1976; Ball and Kirwan, 1977) or examined the effects of market segmentation in their analysis (Nelson, 1975; Straszheim, 1975).

Analyses of market segmentation have met with mixed results. Nelson (1975), for example, found no evidence of urban-suburban differences in an analysis of Washington, D.C. Similarly Ball and Kirwan (1977) were unable to identify spatial submarkets in Bristol based on housing types.

Straszhein (1975), however, identified market differences defined on the basis of race, while Brummell (1981) found differences between urban and suburban areas in an analysis of Thunder Bay. In the latter study,

however, it was unclear whether the differences reflected market differences or inadequacies in specifying the model. Finally, the most comprehensive analysis of submarkets has been provided by Schnare and Struyck (1976). They analyzed the market for single family homes in the suburbs of Boston. Submarkets were defined on three important dimensions: 1) number of rooms, 2) neighbourhood differences as measured by income and 3) location as measured by inner and outer suburbs. Significant differences between submarkets were found supporting the basic concept.

The analysis showed, however, that the submarket model improved the level of explanation only slightly indicating that submarkets seem to account for little of the variation in selling prices. The conclusion reached was that for some purposes at least, the single market model provides a simple and relatively accurate description of house prices. It is important to note, however, that only data from suburban areas were used: no older urban areas were included limiting the variation in attribute mix.

The present study attempts to extend these analyses in two ways. First a comprehensive data set has been developed with an expanded number of structural and neighbourhood variables. For the latter particularly, a number of both social (e.g., income, education etc.) and physical (e.g., % industrial, % multiple family etc.) attributes are included. Second, an expanded number of "submarkets" are examined. Market segments are analyzed on the basis of rooms, income, housing, age, location and price.

III THE SUBMARKET MODEL

The basic hypothesis underlying studies of the determinants of house prices is that the price of a residence depends on the set of attributes associated with that residence. These attributes may be partitioned into three sets: a) structural attributes of the dwelling and lot, \underline{X} , b) social and physical attributes of the neighbourhood, \underline{V} , and c) accessibility attributes of the location, \underline{S} . Incorporating the submarket hypothesis, the basic model may be expressed as:

$$R_{kt} = f_{kt} (\underline{X}, \underline{V}, \underline{S})$$

where R_{kt} refers to the selling price of a residence in submarket k at time t . The analytical problem is to determine if the function, f_{kt} , i.e., set of coefficients reflecting implicit prices, differs across submarkets.

Given a basis for defining the K submarkets and method of estimating the parameters of f_{kt} , i.e., $[B_{ky}]$ then the statistical problem is to test the null hypotheses.

$$B_{kt} = B_{lt}$$

Although time has been included in this general formulation, market changes over time are not analyzed in this study. Henceforth the superscript t will be dropped with the understanding that the analysis refers to a single time period during which it is assumed that no systematic changes in the market has occurred.¹

Given this general conceptual model the analytical problem involves four steps: 1) define (operationally) variables to describe the structural, neighbourhood and locational attributes of residences; 2) develop a set of hypotheses concerning potential submarkets as a basis 3) for partitioning the data set; 4) develop estimates of the parameters in the price-attribute relationship for each data subset as a basis for comparing the nature and extent of submarket differences. In these comparisons attention is focused on differences in the magnitude of specific attribute coefficients as well as the overall pattern of coefficient differences. The extent of submarket differences will be reflected in the magnitude of differences in the overall set of coefficients. The nature of submarket differences will be reflected in those specific coefficients which show the greatest differentiation.

¹This is not strictly true. The data cover a period of six months, June to December, 1980, during which some fluctuation of interest rates occurred.

An important consideration in this analysis is the form of the model specified. In this report two model forms are estimated: a linear model and log-linear model. The general form of these models is as follows:

$$\text{a) Linear: } R = b_0 + \sum b_i X_i$$

$$\text{b) Log-Linear: } R = \exp (b_0 + \sum b_i X_i)$$

R refers to the house price, $[X_i]$ refers to the set of attributes including structural neighbourhood and locational attributes, and $[b_i]$ refers to the set of coefficients to be estimated. The linear form has the advantage of simplicity. The coefficients can be interpreted as the implicit prices of the relevant attributes. This form, however, implies that housing attributes are in some sense divisible. Each room, for example, adds a constant amount to the value of a house. Intuitively, however, rooms are tied together so that the value of an additional room may be related to the number already present. This suggests a non-linear relationship. The log-linear model can accommodate this type of relationship. It also has the advantage that the magnitude of the standard error in the regression analysis varies with the market price; in fact, it is a constant percentage rather than a constant magnitude. From a theoretical viewpoint therefore, the log-linear model is preferable. From a practical viewpoint, however, the "meaning" of the coefficients are less transparent and easy to work with. As a result both models are estimated in the analysis.

IV DATA AND DEFINITION OF VARIABLES

The model specifies three types of variables: structural, neighbourhood, and locational. In each case the number of potential variables which could be included in operationalizing the model is large. Grether and Mieszkowski (1974), for example, used thirty-eight structural variables. In the present case a rich data base allows a relatively large number of variables to be defined. Originally a total of 33 distinct independent variables was defined. Since a number of these were represented by two or more dummy variables more than 40 independent variables were initially included. Subsequent analysis, however, resulted in the removal of a number of variables so that the present analysis is based on 29 variables as listed in Table 1. The only major variable for which no data could be obtained was that for age of the dwelling. Instead a neighbourhood housing age variable is included but this is not considered a good proxy.

The data consist of a sample of 635 single family owner occupied residences sold in the City of Hamilton between June and December 1980. This represents about a 30% sample of all single family houses sold through the MLS service in that period. The majority of dwellings were single-family-detached, but about 8% were semi-detached. No condominiums or apartments were included.

TABLE 1
DESCRIPTION OF VARIABLES - TOTAL MARKET

Variable Description	Mean	Standard Deviation
House Selling Price (\$)	43646.13	17447.65
<u>Structural</u>		
Rooms	5.84	1.12
Bathrooms	1.31	.43
Recreation Room (Yes = 1)	.46	.49
Carpet (Yes = 1)	.40	.49
Fireplace (Yes = 1)	.23	.42
Air Conditioning (Yes = 1)	.06	.24
Brick Exterior (Yes = 1)	.76	.43
Aluminum (Yes = 1)	.09	.29
Driveway (Yes = 1)	.85	.35
Paved Driveway (Yes = 1)	.56	.50
Garage (Single = 1; Double = 2)	.46	.53
Pool (Yes = 1)	.03	.18
Lot Size (100 sq. ft.)	41.94	31.21
Detached (Yes = 1)		
Existing Mortgage (Yes = 1)*	.92	.26
Vendor Mortgage (Yes = 1)*		
<u>Neighbourhood</u>		
Income (Median Household; 1971 in \$ 1980)	19300.62	2917.77
Education (% < Grade 8; 1976)	29.00	11.13
Population Age (% Age 20-34; 1976)	34.08	8.45
Housing Age (% Built Before 1946; 1971)	45.22	35.69
Ownership (% Dwellings Owned)**	67.44	17.23
Multiple Family (% Land Use)**	8.23	7.68
High Rise (% Land Use)**	2.63	4.03
Open Space (% Land Use)**	7.98	12.88
Institutional Open Space (% Land Use)**	5.33	3.68
Industrial (% Land Use)**	5.27	11.14
Retail & Commerical (% Land Use)**	5.47	7.16
<u>Locational</u>		
Distance to CBD (1000 ft.)	18.29	8.45
Distance to Regional Employment (1000 ft.)	23.67	11.47

* Recorded only if the existing or vendor mortgage were 2% below the prevailing interest rate and had more than one year to run.

** Land use data compiled in 1979.

Each observation included not only the selling price but also all the information typically listed on real estate sales descriptions. The observations were obtained from the Metropolitan Hamilton Real Estate Board and were almost exclusively houses sold through the multiple listing service. This source provided the basis for the structural variables in the analysis.

The observations were then located on a map to enable each observation to be related to a specific set of neighbourhood and locational characteristics. Three locational variables were initially defined: distance to the central business district (CBD at King and James); distance to the major regional employment centre (i.e., industrial centre along waterfront); and distance to the nearest local employment centre (i.e., any planning neighbourhood with more than 1,000 employment). The distance to local employment variable was subsequently dropped from the analysis.

Neighbourhood variables were defined using two sources. Land use measures were obtained from recent estimates provided by the City of Hamilton Planning Department. These were relatively fine-grained estimates covering 77 planning neighbourhoods in the city. Other neighbourhood measures were obtained from published census material. The most recent data in many cases were from the 1971 census. As a result the neighbourhood variables, particularly the social variables, represent the pattern of neighbourhood characteristics as of 1971 and to the extent relative changes have occurred

may not be consistent with the pattern as of 1980. While this cannot be tested, the overall pattern has probably changed very little during that period. To facilitate the discussion, the variable median household income has been inflated to 1980 prices using the GNE deflator (2.052 times 1971 income). This, of course does not affect the distribution pattern, only the absolute value.

As noted above the original set of variables identified was somewhat longer than those listed in Table 1. Variables were removed, to simplify the presentation and analysis, after considerable experimentation with different variable combinations and potential submarket divisions. Variables which were consistently insignificant were removed. Typically these variables were highly correlated with an existing variable indicating redundant information. For example, average number of persons per household was correlated with population age (% population aged 20-34). Other variables, however, simply were not important in differentiating housing prices. A variable distinguishing the type of heating, for example, was insignificant in all analyses.

One vexing variable which was finally removed despite a high statistical significance was "private driveway". This variable was unique in that as a structural variable it was relatively highly correlated with a number of neighbourhood variables. As a result it was uncertain whether this variable was serving as a distinct structural characteristic or

quasi-neighbourhood characteristic. Removing this variable improved the interpretability of the results with virtually no loss in overall statistical explanation (R^2).

Finally several variables have not been removed despite consistently low levels of explanation. These variables are either of intrinsic interest (e.g., ownership, type of mortgage) or specific interest with respect to potential submarkets. The land use variables "open space" and "retail space" as examples are not statistically significant when treating the market as a single unitary market, but are of theoretical importance in comparing urban and suburban areas. It is important not to preclude variables which may have important submarket differences.

V. EMPIRICAL ANALYSIS

The purpose of the analysis is to determine the extent to which urban housing markets are segmented into a number of quasi-independent submarkets. The statistical analysis involves two steps. First the parameters of the price-attribute relationship are estimated, via ordinary least squares regression, for the entire sample and each of the subsamples corresponding to hypothesized submarkets. Then the estimated sets of parameters are compared across samples to determine if significant differences occur. If coefficient estimates vary across subsamples then this provides evidence that the relevant market segments are experiencing different market conditions and constitute different submarkets.

The following section examines the price - attribute relationship under the assumption of a single unitary market. This provides the basis for the analysis of submarket differences which follows. In each analysis two forms of the model, linear and log-linear, are presented.

V.1 THE SINGLE MARKET MODEL

In the single market model, the pooled sample is used to estimate the parameters of the house - attribute relationship. The matrix of correlation coefficients is presented in Table 2. The parameter estimates are presented in Table 3. Two specifications for each of the linear and log-linear forms of the model are presented. Specification A includes all independent variables while specification B explicitly excludes the locational variables "distance to the CBD" (DCBD) and "distance to the major regional employment area" (DREMP). The latter refers to the major steel and associated heavy industry district in Hamilton along the waterfront.

We shall examine each component set of variables in turn, i.e., locational, neighbourhood and structural, before comparing the overall results for the linear and log-linear forms. We shall then briefly consider other specifications of the model.

Price	Rooms	Bath	R.R.	Carpet	F.P.	A.C.	Brick	Alum.	Dr.	Paved	Garage	Pool	Lot	Det.	Exist.	Vend.	Income	Educ.	Pop. Age	Hous. Age	Dwn.	Mul.	Fam.	H.R.	Op.Sp.	Inst.	Indus	Retail
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	

1. Rooms	.33																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
----------	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

TABLE 3 SINGLE MARKET COEFFICIENT ESTIMATES

VARIABLE	LINEAR (A)	MODEL (B)	LOG-LINEAR (A)	MODEL (B)
Rooms	3039.0**	3042.8**	0.0513**	0.0524**
Bathrooms	4631.4**	4618.0**	0.0762**	0.0761**
Rec. Room	4437.2**	4532.7**	0.1044**	0.1055**
Carpet	2048.6**	2024.9**	0.0547**	0.0541**
Fireplace	4503.3**	4547.0**	0.1029**	0.1037**
Air Conditioning	4502.2**	4366.5**	0.0732*	0.0688*
Brick	6585.7**	6649.9*	0.2235**	0.2242**
Aluminum	1543.3	1473.8	0.0903**	0.0870**
Driveway	2303.8*	2330.4*	0.1785**	0.1764**
Paved Driveway	1856.0*	1780.3*	0.0628**	0.0602**
Garage	4363.0**	4286.7**	0.0787**	0.0775**
Pool	686.4	925.7	0.0252	0.0319
Lot Size	43.4**	43.7**	0.0008**	0.0008**
Detached	10344.6**	10276.3**	0.2768**	0.2748**
Existing Mort.	864.8	931.0	0.0394*	0.0388*
Vendor Mort.	763.0	1048.5	0.0126	0.0213
Income	1.2**	1.2**	0.00002**	0.00002**
Education	-43.9	-100.3	-0.0032*	-0.0048**
Pop. Age	227.5**	262.3**	0.0043**	0.0047**
Housing Age	-18.2	-36.1*	-0.0006	-0.0010*
Ownership	-42.9	-27.1	-0.0007	-0.0004
% Multi-Fam.	-144.1*	-141.6*	-0.0028*	-0.0025*
% Highrise	-192.2	-199.8*	-0.0035	-0.0036
% Open Space	13.3	24.7	0.0004	0.0007
% Institutional	325.2**	334.4**	0.0052*	0.0058**
% Industrial	-126.4**	-134.9**	-0.0052**	-0.0053**
% Retail	-14.7	-32.5	-0.0005	-0.0009
Dist. CBD	16.5	-	-0.0007	-
Dist. R. Empl	118.8	-	0.0028	-
(Constant)	-35029.8**	-30538.0**	9.0258**	9.1438**
R ²	.738	.737	.780	.778
R ² (Adjusted)	.725	.725	.769	.758
St. Error	9142.9	9147.6	.2077	.2080

** Significant at .01 level

* Significant at .05 level

Locational Variables

Examination of specification A for both model forms indicates that the locational variables are not significant and, except for DCBD in the log-linear form, have the incorrect sign. This raises the question, is location an important factor in determining house prices? Should these variables be included in the analysis?

Locational variables, particularly distance to the CBD, have a long tradition in studies of urban spatial structure (e.g., Mills, 1969; Evans, 1973) and have been included in most empirical studies of house prices (Ball, 1973; Smith, 1977). In some studies these variables have been the main focus of interest (Richardson, Vipond and Furbey, 1974; Jackson, 1979).

Theoretically the distance variables represent measures of accessibility. As such, the coefficients should be negative and significant. That they are not requires some explanation. Three possibilities exist: 1) accessibility is not an important factor influencing house values; 2) accessibility is important but these operational variables are inadequate measures of accessibility; or 3) the effects of accessibility are captured by other variables in the analysis. The first argument needs no clarification. The second argument focuses on the meaning and measurement of accessibility. And the third argument focuses on the relationship between accessibility as measured by these distance variables and other "independent" variables in the analysis.

The meaning and measurement of accessibility is concerned with the way individuals value accessibility and the relationship with distance. For example, the value of accessibility may not be a continuous decreasing function of distance from a centre but a series of discrete steps. In a similar vein, accessibility may be important only beyond some minimum threshold. Alternatively accessibility from a single point may be meaningless. Accessibility may be related to a number of locations and functions. This is why two destinations, one focusing specifically on employment, were included. This may be still too simplistic, however. A more complex measure of accessibility as suggested by Jackson (1979) may be required. In any case the approach adopted may be inadequate to account for the complex relationships between accessibility, location and housing prices, but this cannot be explored further here.

The third argument, that the locational variables are not independent of other variables, is more open to analysis. Statistically this is the problem of multicollinearity: Several independent variables may be highly intercorrelated. In the context of urban structure, this is not unexpected. As Ball and Kirwan (1977) emphasize, distance from the CBD is correlated with other social, physical and structural variables largely because the socio-spatial structure reflects the historical evolution of the built-up area from a central core. The correlation matrix (Table 2) indicates that both distance variables are correlated with housing age, income and education. For DREMP particularly, the correlations are very

high. This makes the statistical task of sorting out the effects of accessibility alone very difficult. On the other hand, the locational impacts, although not necessarily the accessibility impacts, are effectively captured by other variables and may be dropped from the analysis. Comparison of specifications A and B for both linear and log-linear forms confirms that there is very little loss in explanatory power by dropping these variables. As a result, in the following analysis, specification B is used.

Neighbourhood Variables

In both the linear and log-linear forms the importance of neighbourhood variables is clearly demonstrated. Of the 11 variables at least 7 are significant and except for ownership which is not significant in either case, all variables have the expected sign.

The most important finding is that both social and physical characteristics are significant in determining house values. The major social characteristics are income, education and population age. Notice that education refers to the proportion of the population with less than grade eight education: hence the sign of the coefficient is negative. Of these variables, population age and education have seldom been emphasized previously. These variables were included to capture the potential effects of demand segmentation. The age variable, for example, is defined as the

percentage of the population aged 20 to 34 (in 1976). The analysis indicates that as the proportion of the population in this age group increases, house prices increase. This, of course, has been a rapidly expanding age group as the baby boom generation has been entering this age range in increasing numbers. More important than the numbers, however is that this age group has distinct housing needs. This is the life cycle stage where households are acquiring their first house and young families are growing. This distinct group, therefore, may be expected to exert a significant impact on the market as their increasing numbers leads to competition for housing appropriate to their needs. As expected, the presence of this group in an area exerts a positive influence on price. Unfortunately, the data do not allow us to relate this specific age group to any specific housing type which would confirm the concept of market segmentation and support the contention that such segmentation differentially influences prices.

The major physical characteristics of the neighbourhood which are significant are: % land use industrial, % institutional open space, % of land in multi-family units, and % of housing units built prior to 1946 (i.e. housing age). The % of land in highrise use is also significant in the linear form of the model. These findings emphasize the importance of nearby land uses on house prices. Most important are % industrial and % institutional open space. The latter includes parks and school open space which makes this an important social as well as physical characteristic.

In contrast the variable open space is not significant. This variable refers largely to underdeveloped land. The variable is maintained in the analysis to allow for the possibility that this variable may be significant in specific submarkets. For the same reason, the variables ownership and % retail are also maintained in subsequent analyses.

Structural Variables

The importance of structural variables in determining house values is overwhelming.¹ In the log-linear form 14 of the 16 variables are significant at the .05 level. In the linear model form, 12 of the 16 variables are significant. In both cases all coefficients are positive as expected. Indeed, the results are "so expected" that there seems little to discuss. Three observations, however, may be noted to aid subsequent discussions.

First, the magnitude of the coefficients seem reasonable. This is most readily apparent with the linear model. For example, a garage is estimated to add approximately \$4300 to the value of a property; a recreation room adds \$4500; carpeting adds about \$2000; a paved driveway adds about \$1800,

¹ More properly, in regression the independent variables do not determine the house value (the dependent variable) but account for differences in house values.

etc. These values would seem to be reasonable in light of actual costs to install such features. Similar estimates can be generated from the log-linear model. The result is that the estimates "make sense" and lend credibiity to the formulation.

Second, different types of structural variables may be identified.

Essentially 4 types of variables may be observed: 1) basic site variables include number of rooms, lot size and attached or detached structure; 2) internal amenities include number of bathrooms, presence of a recreation room, carpeting, fireplace and central air conditioner; 3) external features include the presence of a pool, garage or driveway, whether the driveway is paved and the type of exterior - brick, aluminum or other; and 4) type of mortgage includes whether an existing mortgage or vendor mortgage at reduced rates, was used. Except for the type of mortgage, which is a feature of the transaction rather than the physical characteristics of the property, and the presence of a pool, all types of structural features are important. In considering submarkets, however, how these different components are valued may be significant.

Finally, the third observation is that the type of mortgage has only a marginal impact on the overall price. In the linear model neither variable is significant. In the log-linear model the presence of an existing mortgage seems to have some impact. In considering submarkets

this feature may be more important for some types of housing than for others. It is also the case that the type of mortgage may be more important than revealed here since the impact of an existing or vendor mortgage may influence whether or not a property is sold at all rather than the final price. This, of course, cannot be determined in this type of analysis.

Comparison of Model Forms

Both the linear and log-linear forms provide excellent accounts of the data. The log-linear model, however, is somewhat superior both in the level of explanation (77.8% vs 73.7%) and in the number of significant variables. Comparing specification B, two variables - aluminum, and existing mortgage - are significant in the log-linear form but not in the linear form. Conversely, % highrise is significant in the linear but not the log-linear form. The pattern of signs is identical in the two cases. Empirically and theoretically, the log-linear form is somewhat superior, but the differences are small. Both forms are used in subsequent analyses.

TABLE 4 SINGLE MARKET COEFFICIENT ESTIMATES: SINGLE DETACHED (N = 587)

VARIABLE	LINEAR (A)	MODEL (B)	LOG-LINEAR MODEL (A)	LOG-LINEAR MODEL (B)
Rooms	2997.8**	2990.2**	0.0467**	.0479**
Bathrooms	4713.8**	4687.4**	0.0754**	.0757**
Rec. Room	4637.5**	4729.1**	0.1137**	.1133**
Carpet	2053.5**	2031.0**	0.0512**	.0510**
Fireplace	4413.7**	4444.7**	0.0959**	.0960**
Air Conditioning	4512.6**	4397.4**	0.0705*	.0679*
Brick	6583.8**	6647.8**	0.2295**	.2288**
Aluminum	1383.2	1331.8	0.0875**	.0838**
Driveway	2248.7	2281.7	0.1741**	.1707**
Paved Driveway	1843.8*	1779.0*	0.0549**	.0528**
Garage	4358.0**	4281.8**	0.0774**	.0770**
Pool	503.8	734.1	0.0325	.0373
Lot Size	43.9**	44.1**	0.0009**	.0009**
Existing Mort.	914.3	982.7	0.0411*	.0385*
Vendor Mort.	808.4	1064.5	0.0193	.0267
Income	1.2**	1.2**	.00002**	.00002*
Education	-63.2	-116.2*	-.0042**	-.0055**
Pop. Age	237.3**	270.9	.0045**	.0044**
Housing Age	-16.2	-32.6	-.0006	-.0007
Ownership	-30.2	-16.6	-.0003	-.0002
% Multi-Fam.	-139.8*	-139.2*	-.0025	-.0020
% Highrise	-130.7	-141.6	-.0016	-.0018
% Open Space	10.8	21.3	.0002	.0005
% Institutional	360.8**	366.6**	.0058*	.0064**
% Industrial	-122.7**	-130.9**	-.0054**	-.0054**
% Retail	-8.5	-25.1	-.0008	-.0010
Dist. CBD	18.4		-.0015	
DREMP	108.6		.0017	
(Constant)	-25680.9**	21403.9**	9.3811**	9.4648**
R ²	.731	.730	.778	.777
R ² (Adjusted)	.717	.717	.767	.767
St. Error	9361.8	9361.1	.2053	.2055

Other Specifications

A variety of other specifications could be estimated from these data by varying the form or the specific variables included. Other forms are generally more complex requiring prior manipulation (e.g., Mark, 1977) or transformation of the data (e.g., Zarembka, 1974) or an alternate estimation procedure (e.g., ridge regression). These offer advantages in overcoming specific data problems such as multicollinearity and merit further analysis. They are not considered here.

Two other specifications have been examined, however, which may be noted. These differ in the specific variables included. One specification excludes the variable "detached". Rather than ignore the variable, however, all attached houses were excluded from the observations and the model recalibrated. The results are shown in Table 4. The purpose was to determine if including this subsample distorted the coefficient estimates, i.e., attached and detached houses form separate submarkets. Comparing Tables 3 and 4 indicates that the coefficient differences are very small; the pattern of signs is identical; the R^2 and standard error values are virtually identical. Including attached houses in the sample does not seem to distort the results. Consequently, all observations are included in subsequent analyses.

TABLE 5 SIMPLIFIED SINGLE MARKET COEFFICIENT ESTIMATES

VARIABLE	LINEAR (A)	MODEL (B)	LOG-LINEAR MODEL (A)	LOG-LINEAR MODEL (B)
Rooms	2965.7**	3090.4**	.0529**	.0534**
Bathrooms	4670.4**	4579.7**	.0761**	.0745**
Rec. Room	4529.0**	4568.7**	.1036**	.1055**
Carpet	2057.5**	2010.0**	.0533**	.0540**
Fireplace	4466.9**	4522.3**	.1037**	.1036**
Air Conditioning	4540.7**	4609.5**	.0752*	.0739*
Brick	6573.5**	6110.3**	.2248**	.2244**
Aluminum	1644.2		.0883**	.0851**
Driveway	2450.8*	2513.3*	.1761**	.1765**
Paved Driveway	1985.1*	1845.9*	.0649**	.0623**
Garage	4343.8**	4272.2**	.0781**	.0778**
Pool				
Lot Size	44.3**	42.8**	.0008**	.0008**
Detached	10292.4**	10276.8**	.2744**	.2744**
Existing Mort.	808.7	835.7	.0354*	.0373*
Vendor Mort.				
Income	1.4**	1.0**	.00002**	.00002**
Education		-122.1*	-.0036*	-.0052**
Pop. Age	238.8**	268.1**	.0042**	.0048**
Housing Age		-43.5	-.0006	-.0011**
Ownership	-60.4*1			
% Multi-Fam.	-169.1**	-126.1*	-.0026*	-.0022
% Highrise	-193.7*	-183.8*	-.0024	-.0034
% Open Space		40.5		.0010
% Institutional	-318.0**	362.9**	.0060**	.0063**
% Industrial	-131.5**	-129.5**	-.0051**	-.0051**
% Retail				
Dist. CBD				
DREMP	158.1**1		.0032*1	
(Constant)	-39015.8**	-29138.1**	9.0189**	9.1493**
R ²	.738	.736	.779	.778
R ² (Adjusted)	.728	.726	.770	.769
St. Error	9104.0	9126.8	.2071	.2075

The second alternative specification allows the data a larger role in determining the variables which are included. In Table 3 all variables are entered even though they may have an extremely low level of significance. In the present case only variables with t-values greater than 1.0 are included¹. This minimizes the standard error of estimate although the R^2 values are marginally lower. The results are shown in Table 5. The advantage of this "best fit" approach is that highly insignificant variables are removed so that the presentation of results is much clearer. In comparing Tables 5 and 3 the major difference is that in specification A, for both model forms, "distance to regional employment" is now significant using a 2-tail test. Theoretically, however, the sign is in the wrong direction. We shall continue to focus on specification B.

This best fit approach, along with the "all variable approach" is used in analysing the submarket models below. This approach is appropriate in that we do not want to force variables into the analysis which are not significant in a specific submarket. It does, however tend to increase the probability of rejecting the single market model and is thus not a conservative approach. The "all variables" approach, in contrast, appears overly conservative. Both results are included.

¹ The actual criterion is that the last variable entered in the stepwise regression has a t-value of 1.0.

V.2 SUBMARKET MODELS

The purpose of the analysis is to determine the extent that submarkets exist in urban housing. The basic hypothesis is that the urban housing market is segmented into a number of quasi-independent submarkets. This implies that at any given moment in time, differences in house prices may reflect not only differences in housing attributes but also differences in submarket conditions. Differences in market conditions are reflected in the parameters of the price-attribute relationship. The model implies, therefore, that if submarkets exist and conditions vary, then the magnitude of the coefficients estimated in the price-attribute relationship should vary across submarkets. Conversely, if differences in the magnitude of the estimated coefficients can be demonstrated then this provides evidence of the existence of submarkets.

This logic provides the basis for the submarket analysis. The total sample is partitioned into subsamples corresponding to the hypothesized submarkets. The price-attribute relationship is then estimated for such submarket separately and the coefficients compared across submarkets. Both the "all variables", (AV) and "best fit" (BF) solutions are provided permitting two comparisons of the results. Submarkets are defined along five dimensions: house size (number of rooms), neighbourhood housing age,

median income, location and house price. Statistical descriptions of the submarket samples are provided in Table 6. The coefficient estimates for each of the hypothesized submarkets are presented in Tables 7 through 10 for the linear and log-linear models as indicated.

House Size

House size is a fundamental aspect of housing influencing its price. Large houses are more expensive, on average than small houses. Is this difference the result only of size or do large and small houses form separate markets? In this analysis large houses are defined as those with more than five rooms: small houses have five rooms or less. Using this division there are 409 large houses and 226 small houses in the sample. As shown in Table 6 the average price of small houses is \$37,983: the average price of large houses \$46,775. The average number of rooms for small houses is 4.7 and for large houses 6.5. Other than these two key variables few other variables show large differences. The only major differences are in the proportion of houses with brick exteriors and with garages. Other than size, therefore, large and small houses show few structural, neighbourhood or locational differences.

TABLE 6
DESCRIPTION OF "SUBMARKETS": MEAN VALUES

	Total Sample	HOUSE SIZE		NEIGHBOURHOOD AGE		MEDIAN INCOME		URBAN/SUBURBAN		REGIONS			HOUSE PRICE	
		Large GT 5 Rms	Small LE 5 Rms	New Le 10% Pre'46	Old GT 10% Pre'46	High GT \$20000	Low LE \$20000	Urban	Suburban	Southwest	Northwest	Mountain	Low LE \$40000	High GT 40000
Selling Price	43646.13	46774.91	37983.85	55306.18	39383.31	54983.60	35734.14	33890.50	52846.33	44827.78	35891.34	52772.74	27887.83	54932.47
Rooms	5.84	6.47	4.70	5.82	5.85	5.92	5.78	5.90	5.79	6.07	5.80	5.79	5.63	5.99
Bathrooms	1.31	1.41	1.14	1.42	1.27	1.41	1.25	1.24	1.39	1.36	1.25	1.37	1.18	1.41
Rec. Rooms	.46	.47	.43	.76	.34	.68	.30	.22	.68	.39	.30	.68	.15	.68
Carpet	.40	.40	.42	.50	.37	.49	.35	.32	.48	.37	.35	.50	.26	.51
Fireplace	.23	.29	.12	.29	.21	.33	.16	.16	.29	.31	.15	.29	.06	.36
AC	.06	.07	.05	.11	.05	.10	.04	.05	.08	.05	.06	.07	.03	.08
Brick	.76	.83	.62	.91	.70	.89	.66	.67	.37	.77	.47	.85	.52	.92
Aluminum	.09	.05	.16	.06	.11	.06	.12	.11	.27	.08	.31	.08	.16	.05
Driveway	.85	.85	.85	.99	.80	.98	.76	.71	.98	.88	.74	.97	.67	.98
Paved	.56	.56	.57	.82	.47	.74	.44	.39	.73	.55	.45	.72	.32	.74
Garage	.46	.50	.37	.48	.45	.52	.41	.40	.51	.46	.43	.50	.32	.56
Pool	.03	.04	.03	.07	.02	.07	.01	.01	.06	.03	.01	.07	.00	.06
Lot Size	41.94	43.24	39.60	50.69	38.75	52.99	34.24	31.98	50.83	36.12	34.59	53.32	32.86	48.45
Detached	.92	.90	.96	.91	.93	.90	.94	.93	.91	.95	.93	.91	.92	.93
Existing Mort.	.39	.41	.36	.51	.35	.51	.31	.32	.50	.25	.48	.50	.29	.46
Vendor Mort.	.05	.05	.04	.01	.06	.04	.05	.03	.24	.16	.15	.02	.03	.06
Income	19300.62	19436.16	19055.32	22612.59	18089.79	22191.52	17283.17	17029.09	21392.98	18504.56	17656.87	21688.32	17328.77	20712.89
Education	29.00	28.29	30.29	21.98	31.57	20.85	34.69	37.99	20.88	24.44	37.79	20.84	36.76	23.45
Pop. Age	34.08	35.72	31.13	34.19	34.04	36.05	32.71	32.48	35.60	36.22	32.65	34.82	32.02	35.56
Housing Age	45.22	47.71	40.71	3.58	60.44	18.38	63.94	74.79	17.60	60.80	65.50	12.37	66.98	29.63
Ownership	67.44	66.45	69.22	68.56	67.03	74.93	62.21	63.85	70.72	58.56	65.73	73.86	65.85	68.58
% Multi-Fam.	8.23	9.77	5.45	6.31	8.93	6.77	9.24	10.78	5.99	6.52	10.22	6.81	9.14	7.58
% Highrise	2.63	2.77	2.37	3.61	2.27	2.32	2.85	2.65	2.67	2.74	2.78	2.46	2.34	2.84
% Open Space	7.98	8.52	7.00	7.42	8.18	8.64	7.52	7.05	8.94	17.98	7.18	4.19	6.91	8.74
% Institutional	5.33	5.26	5.47	6.47	4.92	5.94	4.90	5.15	5.58	5.41	4.87	5.97	4.90	5.64
% Industrial	5.27	5.50	4.85	.45	7.03	.33	8.71	10.30	.78	7.03	8.86	.25	9.49	2.25
% Retail	5.47	4.60	7.06	3.37	6.24	2.76	7.37	8.13	3.12	2.85	8.30	3.49	8.01	3.65
Dist. CBD	18.29	17.75	19.27	24.02	16.20	22.22	15.55	14.65	21.20	11.10	18.37	21.14	15.57	20.24
DREMP	23.67	24.46	22.24	31.91	20.66	32.26	17.68	13.62	32.86	28.80	13.16	33.75	16.12	29.08
No Observ.	635	409	226	170	465	261	374	303	332	117	282	236	265	370

1. For detailed description of variables see Table 1.

The results of the regression analysis for large and small houses are presented in columns (2) and (3) in Tables 7 through 10. The analytical question is: do the coefficient estimates differ between the two potential submarkets? Focusing on the BF model (Tables 7 and 8) a number of differences are revealed. Relatively large differences are evident among the basic site variables, rooms and lot size; among internal amenities, bathrooms, recreation room, carpet and air conditioning; among external features, brick exterior, driveway, paved driveway, garage and pool; and among neighbourhood characteristics, education level, % highrise and % open space. The size and number of these differences suggest that large and small houses may form separate markets. However, are these differences statistically significant?

To test whether the pattern of coefficients are significantly different across submarkets a Chow test may be performed (Silk, 1979; Smit, 1979). The null hypothesis is that there are no differences between coefficients. The alternative hypothesis is that at least one coefficient differs across submarkets. Essentially, the Chow test is a comparison of the submarket and single market models. The test examines the difference between the residual (or unexplained) sum of squares associated with the submarket model, which allows the coefficients to differ between submarkets, and the residual sum of squares associated with the restricted single market

TABLE 7
BEST FIT LINEAR MODEL: COEFFICIENT ESTIMATES

Variable	HOUSE SIZE			NEIGHBOURHOOD AGE		MEDIAN INCOME		URBAN/SUBURBAN		REGIONS			HOUSE PRICE	
	Single Market	Large GT 5 Rms	Small LE 5 Rms	New Le 10% Pre '46	Old GT 10% Pre '46	High GT \$20000	Low LE \$20000	Urban	Suburban	Southwest	Northwest	Mountain	Low LE \$40000	High GT 40000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Rooms	3090.4**	4131.9**	3277.6**	4249.5**	2738.4**	5440.6**	2203.1**	2330.5**	4192.7**	2876.0**	2214.3**	4575.8**	251.7	5269.3**
Bathrooms	4579.7**	4756.2**		6571.6**	3204.2**	5823.2**	2231.8*		6700.9**		4966.9**	5628.9**	2513.9**	5144.0**
Rec. Rooms	4568.7**	4227.1**	5969.7**	5585.8**	4637.9**	4026.8**	5238.7**	4862.1**	4628.1**	2347.9	5809.0**	5181.4**	1573.6*	2211.4*
Carpet	2010.0**	1246.7	3370.6**	2793.7*	1533.9	1792.3	2576.9**		2786.8**	2518.9	2562.5*	1750.0	2863.6**	1304.6
Fireplace	4522.3**	4548.5**	4103.9**		6957.5**	2444.4*	7213.5**	8920.2**	2201.3*	5854.9**	5858.3**	2656.4*		1690.8
AC	4609.5**	5611.5**	1776.7	2159.0	6571.5**	4756.3*		4412.2*	4401.5**	9378.1**		4341.4*	2701.5	4376.3*
Brick	6110.3**	7183.9**	5347.3**	6811.9	6788.7**	8088.8**	5860.7**	6089.2**	7358.0**	7110.6**	5160.9**	9348.8**	3574.5**	6268.3*
Aluminum					1870.9		2121.6	2571.7		3356.4		3441.5	1908.9*	3683.6
Driveway	2513.3*		5030.0**		2740.5*		3176.7**	4433.8**		5188.8*	3043.6*		2386.3**	
Paved	1845.9*	1527.9	2945.7**		1999.6*		1901.5*	3130.3**		1506.3	2834.7*		2978.1**	
Garage	4272.2**	5513.2**	1811.2*	6778.2**	2680.7**	5471.3**	2807.3**		5749.5**	1796.1	2266.8*	5975.4**	1354.9*	4736.7**
Pool			5765.4**		3853.3		4334.6			5805.8				
Lot Size	42.8**	50.0**	31.7*	152.8**	24.4*	44.8**	33.7*	47.0**	52.8**	245.5**	46.3*	27.6*	18.1*	45.4
Detached	10276.8**	9933.9**	8941.4**	9946.4**	9762.1**	10783.4**	9212.9**	9469.9**	9815.5**	7043.3**	10979.3**	8788.1**	6556.5**	9623.6**
Existing Mort.	835.7	1238.5	886.9				2798.8**	2405.5*		2058.9	2448.4*		1196.7*	
Vendor Mort.			2602.2			3363.2		2663.5	3799.7*					
Income	1.0**	0.9**	1.0**	1.6*	1.5**	0.7	1.5**	1.1**	1.2**		1.7*	1.1**	.7**	.7*
Education	-122.2**	-188.3*			-94.2		-41.9	-241.4**	288.1	-428.6**	19.9	659.1*	-93.6*	130.6
Pop. Age	268.1**	253.9**	273.9**	176.5	262.9**	188.9*	350.6**		296.3**	-732.7**	376.7*	458.6**	92.4	249.6**
Housing Age	-43.5	-56.0	-37.8*		-38.7*	-41.4	-61.0**		-32.3	-142.1				-44.5*
Ownership				-156.9		-74.6			-58.5		-117.3	81.6		
% Multi-Fam.	-126.1*	-126.4		-434.3*	-114.0*	-179.4			-227.4*	946.2	-205.2*	-467.9**		-244.9**
% Highrise	-183.8	-196.2	-376.7*	-375.7*		-224.5	-120.2		-274.4*	61.0	-323.9*	-202.6		-186.9*
% Open Space	40.5	76.2*		316.3*			37.2	57.9		80.8		161.8	24.9	
% Institutional	362.9**	361.8**	236.8*	365.2**	331.2*	201.4	654.1**	403.0*	240.6*	666.7		478.3**	412.4**	219.4*
% Industrial	-129.5**	-119.1*	-135.9**		-120.8**		-97.7**	-105.7**		-217.3*	-207.0**	709.9	-74.0**	
% Retail				-302.2*		-295.9			-293.8*		-66.1	-367.4		-203.8*
Constant	-29138.1**	-31607.5**	-27552.9**	-41814.9**	-33715.4**	-26162.9*	-37285.7**	-14099.7*	-42700.1**	30381.1**	-37022.4	-65969.9**	-3155.2	-30168.3**
R ²	.736	.716	.814	.695	.719	.614	.700	.651	.676	.800	.699	.706	.643	.565
R ² (Adjusted)	.726	.701	.796	.661	.706	.584	.681	.631	.655	.749	.675	.676	.615	.542
St. Error	9126.6	10367.4	5684.2	8698.1	8854.7	9852.5	8031.7	8778.8	8714.9	7384.3	9620.4	8354.9	4825.4	8926.7
N	635	409	226	170	465	261	374	303	332	117	232	236	265	370

Combined R²

** Significant at .01 level

* Significant at .05 level

TABLE 8
BEST FIT LOG-LINEAR MODEL: COEFFICIENT ESTIMATES

Variable	HOUSE SIZE			NEIGHBOURHOOD AGE		MEDIAN INCOME		URBAN/SUBURBAN		REGIONS			HOUSE PRICE	
	Single	Large	Small	New	Old	High	Low	Urban	Suburban	Southwest	Northwest	Mountain	Low	High
	Market	GT 5 Rms	LE 5 Rms	Le 10% Pre '46	GT 10% Pre '46	GT \$20000	LE \$20000						LE \$40000	GT 40000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Rooms	.0534**	.0569**	.0746**	.0702**	.0512**	.0901**	.0427**	.0415**	.0680**	.0635**	.0412**	.0739**	.0177	.0748**
Bathrooms	.0745**	.0782**		.0814**	.0773**	.0743**	.0649*	.0592	.0920**		.0916*	.0820**	.0834*	.0701**
Rec. Rooms	.1055**	.0945**	.1343**	.1360**	.1015**	.1078	.1101**	.0996**	.1187**		.1187**	.1352**	.0569	.0472**
Carpet	.0540**	.0271	.0905**	.0474*	.0523*	.0246	.0810	.0379	.0570**	.0694*	.0721*	.0489*	.1081**	.0188
Fireplace	.1036**	.1070**	.1160**		.1515**	.0466*	.1722**	.2150**	.0447*	.1168**	.1286**	.0516*		.0872*
AC	.0739*	.0689	.0783	.0531	.0901*	.0650*		.0756	.0428	.0876	.0767	.0611	.1100	.0681**
Brick	.2244**	.2405**	.1966**	.1786**	.2274**	.1940**	.2046**	.2292**	.2135**	.2183**	.2226**	.2569**	.1535**	.1194**
Aluminum	.0851**	.1068*	.0524		.1026**		.1078**	.1216*	.0850*	.1192	.0707	.1329*	.0814*	-.0699
Driveway	.1765**	.1234**	.2636**	.2678*	.1629		.1578**	.1871**	.0589	.2848**	.1470**		.1048**	-.0699
Paved	.0623**	.0565**	.0892**		.0638**	.0285	.0623*	.0948**			.1044**		.1041**	
Garage	.0778**	.0867**	.0484*	.1020**	.0618**	.0956**	.0732		.1001**	.0444	.0506	.1051**	.0526*	.0825**
Pool			.1103		.0921	.0738*			.0469	.1359	.1460			
Lot Size	.0008**	.0010**		.0021**	.0006*	.0007*	.0008*	.0010*	.0008**	.0062**	.0008	.0007*	.0006	.0007**
Detached	.2744**	.2674**	.2927**	.1909**	.3130**	.2101**	.3291**	.3559**	.1903**	.1795*	.3481**	.1750**	.2985**	.1702**
Existing Mort.	.0373*	.0422*	.0354		.0484*	-.0241		.0951**	-.0187	.0851*	.0785**		.0646*	
Vendor Mort.				.1113		.0586	.1011**		.0579	.0660				
Income	.00002**	.00001*	.00002**	.00003**	.00003**		.00004**	.00003**	.00002**		.00005**	.00003**	.00002*	.00002**
Education	-.0052**	-.0070**			-.0040**	.0054*	-.0025	-.0064**		-.0102**	-.00007	.0096*	-.0023	.0015
Pop. Age	.0048**	.0046**	.0064**	.0029	.0049**	.0019	.0077**		.0049**	-.0221**		.0076**	.0030	.0041**
Housing Age	-.0011**	-.0011*	-.0012*		-.0010	-.0010*	-.0012		-.0010*	-.0059**			-.0008	-.0006*
Ownership				-.0029		-.0012			-.0018* ¹		-.0043** ¹			-.0009
% Multi-Fam.	-.0022	-.0030*		-.0052*	-.0022				-.0030	.0286** ¹	.0049*	-.0062**		-.0039**
% Highrise	-.0034	-.0026	-.0134**	-.0054					-.0052*	.0024	-.0070	-.0047	-.0062	-.0043*
% Open Space	.0010	.0017*	-.0013	.0034			.0012	.0017		.0024		.0028		
% Institutional	.0063**	.0057	.0090*	.0040*	.0108**		.0223**	.0147**		.0222*		.0031	.0175	.0023
% Industrial	-.0051**	-.0050**	-.0069**	.0091	-.0049**	.0168	-.0043**	-.0045**		-.0058*	-.0073**	.0172	-.0047**	
% Retail				-.0053*		-.0042			-.0044*			-.0070*		-.0029
Constant	9.1493**	9.2717**	8.7454**	8.8651	8.8399**	9.8324**	8.3702**	8.9627**	9.2356**	10.4139**	8.7971**	8.6724**	9.1026**	9.5998**
R ²	.778	.765	.814	.727	.751	.656	.733	.705	.715	.823	.761	.714	.621	.592
R ² (Adjusted)	.769	.751	.796	.690	.739	.628	.718	.686	.695	.783	.741	.687	.590	.567
St. Error	.2075	.2221	.1722	.1384	.2231	.1624	.2217	.2396	.1491	.1693	.2320/	.1553	.2044	.1361
N	635	409	226	170	465	261	374	303	332	117	282	236	265	370
Combined R ²		.788		.790		.800		.805		.815			.860	

** Significant at .01 level
* Significant at .05 level
¹ 2 tail test

TABLE 9
ALL VARIABLES LINER MODEL: COEFFICIENT ESTIMATES

Variable	HOUSE SIZE			NEIGHBOURHOOD AGE		MEDIAN INCOME		URBAN/SUBURBAN		REGIONS			HOUSE PRICE	
	Single	Large	Small	New	Old	High	Low	Urban	Suburban	Southwest	Northwest	Mountain	Low	High
	Market	GT 5 Rms	LE 5 Rms	Le 10% Pre'46	GT 10% Pre'46	GT \$20000	LE \$20000						LE \$40000	GT 40000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Rooms	3042.8**	4117.7**	3120.8**	4135.8**	2734.5**	5301.7	2177.41**	2236.8**	4220.3**	3179.9**	2298.4**	4479.9**	360.7	5223.6**
Bathrooms	4618.0	4769.5**	826.7	6878.0**	3206.5**	5988.9**	2319.45**	1186.2	6650.6**	-1470.9	4819.5**	5558.5**	2448.2**	5015.1**
Rec. Rooms	4532.7**	4285.5**	5895.7**	5489.5**	4644.8**	3888.3**	526-.34**	4712.9**	4612.8**	2476.4	5613.4**	5204.7**	1710.2*	2164.0*
Carpet	2024.9**	1392.8	3343.2**	2888.4*	1532.8	1727.1	2494.14**	774.0	2730.5**	2093.4	2495.2*	1767.8	2920.3**	1347.3
Fireplace	4547.0**	4568.9**	4279.4**	223.8	6890.2**	2426.9	7153.76**	8737.6**	2121.3*	5508.0**	5593.6**	2851.4*	928.7	1763.8*
AC	4366.5**	5231.1**	1641.9	2332.6	6701.5**	4782.6*	750.46	4282.9*	4209.5*	9595.9**	2781.9	4412.4*	2543.1	4381.4**
Brick	6649.9*	7796.9**	5286.6	7727.3**	6695.9**	9279.8	5667.13**	5943.2**	8480.1**	7228.7**	5680.8**	9296.0**	3637.3**	6016.7*
Aluminum	1473.8	2248.1	-	756.5	1901.6	2146.1	2154.13	2008.6	2211.2	2671.4	249.6	3129.6	1680.9*	3532.3
Driveway	2330.4*	374.7	5132.6**	8855.0	2672.3*	-1478.2	3180.09**	4631.8**	-511.1	5282.9*	2774.6	873.8	2476.6**	-3784.4
Paved	1780.3*	1332.9	2896.8**	-460.7	2007.5*	673.9	1881.11*	2779.9*	-	1839.2	2819.6*	625.6	2793.4**	290.1
Garage	4286.7**	5475.1**	1766.3*	6809.3**	2745.9**	5482.9**	2797.41**	318.5	5754.2**	1759.9	2641.9*	5613.3**	1343.9*	4895.8**
Pool	925.7	-1379.7	5497.3**	-6664.8	3861.6	1070.3	4309.80	3340.2	864.4	5750.4	4249.5	-428.1	-	-369.7
Lot Size	43.7**	51.3	34.4*	147.9**	25.3*	43.3*	34.12	41.4*	52.3**	245.7**	44.1*	25.7	16.3	48.0
Detached	10276.4**	10016.4**	8723.9**	9880.1**	9927.5**	10530.80**	9338.70**	9187.7**	9843.5**	6832.6*	11155.9**	8887.9**	6345.9**	9598.9**
Existing Mort	931.0	1281.4	829.7	-513.8	899.1	-922.3	2899.38**	2150.5*	-566.2	2233.3	2447.6*	-720.2	1187.0*	431.9
Vendor Mort.	1048.5	403.3	2585.9	4221.8	1080.4	2713.4	1150.73	-3118.8	2182.8	3837.1*	-2386.3	2781.7	-360.5	-
Income	1.2**	1.1**	1.4**	1.7*	1.6**	.5	1.88**	1.9*	1.0**	.3	1.3	1.4**	.6*	.8*
Education	-100.3	-150.9*	53.7	-	-72.9	-111.0	-19.23	-122.19	260.7	-390.2*	-11.8	740.8*	-80.3	140.2
Pop. Age	262.3**	247.7**	243.9	206.7	256.1**	178.2*	297.69**	67.2	306.8**	-535.4	455.9*	499.1**	90.3	234.2**
Housing Age	-36.1*	-44.3	-23.5	-167.2	-43.3*	-39.6	-46.74	15.9	-55.5	-142.5	-19.4	52.3	-13.3	-40.5
Ownership	-27.1	-32.6	-51.7	-150.1	-25.4	-79.6	-51.08	-80.9	-53.5	41.2	-63.8	81.1	-	-37.0
% Multi-Fam.	-141.6*	-156.1*	-10.9	-453.5*	-100.9	-180.1	-30.17	-104.8	-236.6*	982.4	-142.6	-489.0**	-27.9	-242.6**
% Highrise	-199.8*	-202.5	-461.0*	-379.6*	-27.2	-195.4	-142.03	62.9	-308.6*	64.6	-365.7*	-184.4	-82.9	-273.6*
% Open Space	24.7	51.3	-51.1	346.1	19.9	-9.3	30.29	45.4	40.3	67.2	32.3	145.8	13.9	14.4
% Institutional	334.4**	335.1**	217.9	374.7*	292.7*	193.4	667.33**	353.8	257.1*	767.4	113.2	491.4**	427.5**	229.8*
% Industrial	-134.9**	-130.1	-146.3**	511.7	-116.8	451.7	-95.51	-103.2*	-97.0	-245.9	-188.4**	697.3	-81.2**	-71.9
% Retail	-32.5	-99.6	-	-366.2*	19.7	-340.5	30.13	50.9	-279.3*	110.5	-51.3	-378.8*	-	171.2
Constant	-30538.0**	-34464.7**	-31847.0**	-53065.3**	-35275.3**	-19114.4	-40333.31**	30083.0	-39654.1**	14171.2	-36084.4*	-75919.3**	-29.3	-25613.1*
R ²	.737	.718	.815	.700	.721	.617	.701	.656	.678	.802	.701	.708	.645	.568
R ² (Adjusted)	.725	.698	.792	.645	.703	.573	.678	.622	.650	.740	.669	.669	.610	.535
St. Error	9147.6	10416.2	5717.9	8895.1	8895.8	9979.48	8074.46	8886.2	8772.6	7514.7	9709.3	8446.5	4855.3	8983.6
N	635	409	226	170	465	261	374	303	332	117	282	236	265	370
Combined R ²		.752	.762	.762	.766	.763		.766			.775		.827	

Variable	HOUSE SIZE			NEIGHBOURHOOD AGE		MEDIAN INCOME		URBAN/SUBURBAN		REGIONS			HOUSE PRICE	
	Single	Large	Small	New	Old	High	Low	Urban	Suburban	Southwest	Northwest	Mountain	Low	High
	Market	GT 5 Rms	LE 5 Rms	Le 10% Pre '46	GT 10% Pre '46	GT \$20000	LE \$20000						LE \$40000	GT 40000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Rooms	.0524**	.0556**	.0711**	.0689**	.0504**	.0854**	.0443**	.0446**	.0673**	.0700**	.0413**	.0724**	.0197	.0751**
Bedrooms	.0761**	.0792**	.0196	.0815**	.0775**	.0780**	.0656	.0542	.0910**	-.0399	.0908*	.0856**	.0847*	.0696**
Bathrooms	.1055**	.0959**	.1348**	.1303**	.1016**	.1071**	.1107**	.1000**	.1164**	.0378	.1145**	.1319**	.0536	.0469**
Carpet	.0541**	.0310	.0875**	.0465*	.0540*	.0286	.0757**	.0424	.0552**	.0494	.0766*	.0448*	.1082**	.0164
Replace	.1037**	.1067**	.1145**	.0190	.1511**	.0469*	.1690**	.2099**	.0457**	.1186**	.1221	.0566*	.0322*	.0373*
	.0688*	.0654	.0839	.0415	.0920*	.0644*	.0412	.0815	.0484	.1036	.0879	.0598	.1074	.0669**
Deck	.2242**	.2372**	.1967**	.1814**	.2259**	.2247	.2051**	.2286**	.2107**	.2083**	.2262**	.2436**	.1547**	.1191**
Minimum	.0870**	.1074*	.0585	-	.1023**	.0959	.1065**	.1053*	.0817*	.1178	.0650	.1148*	.0812*	.0651
Driveway	.1764**	.1222**	.2516**	.2603*	.1619**	-.0123	.1565**	.1874**	.0514	.2806**	.1477**	.0579	.1056**	-.0685
Bed	.0602**	.0546*	.0876**	.0077	.0635**	.0304	.0643*	.0883**	.0142	.0400	.1081**	.0165	.1029**	.0062
Age	.0775**	.0892**	.0468*	.1039**	.0625**	.0917**	.0718**	.0233	.0981**	.0456	.0537	.0938**	.0524*	.0816**
Lot	.0319	-.0178	.1106	.0173	.0944	.0582	.0772	.1245	.0454	.1552	.1685	.0246	-	-.0125
Size	.0008**	.0011**	.0004	.0021**	.0006*	.0007*	.0008*	.0010*	.0008**	.0058*	.0008	.0006*	.0005	.0007**
Attached	.2748**	.2659**	.2993**	.1875**	.3160**	.2033	.3299**	.3492**	.1931**	.1859**	.3544**	.1818**	.2973**	.1716**
Listing Mort	.0388*	-	.0712	.111	.0275	.0602	.0228	-.0728	.0599	.0701	-.0561	.0705	-	-.0036
Ref. Mort.	.0213	-	.0712	.1111	.0275	.0602	.0228	-.0728	.0599	.0701	-.0561	.0705	-	-.0036
Home	.00002**	.00002*	.00001	.000003**	.00003*	.00001	.00005*	.00004*	.00002**	-.00002	.00004	.00003**	.00001	.00002**
Location	-.0048**	-.0063**	-.0006	-	-.0037*	-.0040	-.0024	-.0040	.0020	-.0131**	-.0009	.0114	-.0030	.0013
Age	.0047**	.0045**	.0066**	.0039	.0049**	.0030*	.0075**	.0014	.0057**	-.0197*	.0053	.0087**	.0035	.0042
Living Age	-.0010	-.0008	-.0014*	-.0049	-.0011*	-.0005	-.0011	.0002	-.0010	-.0051*	-.0002	.0002	-.0009	-.0006
Partnership	-.0004	-.0010	.0012	-.0024	-.0005	-.0018	-.0005	-.0017	-.0015* ¹	.0018	-.0025	.0004	-.0010	-.0008
Multi-Fam.	-.0025*	-.0037*	.0020	-.0049	-.0019	-.0017	-.0003	-.0026	-.0032*	.0286**	-.0041	-.0068	-.0004	-.0038**
Highrise	-.0036	-.0034	-.0111	-.00										

model. The residual sum of squares will always be less in the disaggregated submarket model, but whether this reduction is significant depends on the size of the difference relative to the number of additional parameters estimated. The ratio of the difference to the residual sum of squares of the submarket model is distributed as F. In the AV formulation, the submarket model requires 56 parameter estimates, 28 more than the single market model. In the BF formulation the submarket models requires 43 parameters estimates for the linear model, 20 more than for the "best fit" single market model, and 45 parameters estimates for the log-linear model, 21 more than for the "best fit" single market model.

The Chow test results are shown in Table 11(a). Only for the linear model in the best fit case are the differences statistically significant. This is weak evidence to support the hypothesis that small and large houses represent different submarkets. At least with this submarket definition the hypothesis cannot be accepted.

To gain a more intuitive feel for the results the standard errors and combined R^2 statistics may be used. The combined R^2 statistic is the overall level of explanation achieved by the submarket model combining the subsample results. Even in the best fit case the level of explanation improves only slightly, from .736 to .750 in the linear model and from .778 to .788 in the log-linear model, despite the large increase in the number of parameters estimated. Similarly, the standard errors for the submarket model diverge from the single market model. There is a distinctly better

TABLE 11
TEST OF SUBMARKET DIFFERENCES

	F VALUES	
	<u>BEST FIT</u>	<u>ALL VARIABLES</u>
A) House Size		
Linear	1.69*	1.26 n.s.
Log-linear	1.37 n.s.	1.11 n.s.
B) Neighbourhood Age		
Linear	3.59**	2.21**
Log-linear	1.73**	1.39 n.s.
C) Median Income		
Linear	3.18**	2.27**
Log-linear	3.68**	2.45**
D) Urban/Suburban		
Linear	4.66**	2.58**
Log-linear	4.77**	3.02**
E) Regions		
Linear	2.01**	1.59**
Log-linear	2.98**	2.15**
F) House Price		
Linear	18.11**	12.70**
Log-linear	18.12**	13.58**

** Significant at .01 level

* Significant at .05 level

fit for the small house segment but a worse fit for the large house sample. Thus the single market model does virtually as well as the single market model with many fewer parameter estimates.

Finally, it is important to note a significant difference between the linear and log-linear formulations. In the linear model the coefficient for the variable rooms is larger for large houses than small houses. In the log-linear model the coefficient is larger for small houses. The latter result is consistent with the interpretation that as the number of rooms in a house increase, the value of an additional room, that is the marginal value, decreases. This is consistent with economic theories of individual choice and suggests that a non-linear relationship between rooms and price exists. This may be one reason why the log-linear model provides consistently higher levels of explanation.

Neighbourhood Housing Age

The period during which a neighbourhood was first developed is an important indicator of the type of housing, characteristics and location of that neighbourhood. Older neighbourhoods are typically located closer to downtown with numerous two and three storey brick houses on small lots. Many do not have side driveways or garages; many have been divided into multi-family units or replaced by highrises; open space is scarce; and numerous other commercial and industrial uses are intermixed with

residential uses. In contrast, new neighbourhoods, located near the city edge, are characterized by bungalows or split levels on large lots with paved driveways; there are few multi-family houses and only an occasional high rise cluster; plenty of parks and open space are available with only an occasional corner store or segregated shopping mall. In short, old and new neighbourhoods are typically associated with quite different housing environments. The age of a neighbourhood, therefore, would seem to be an important variable capturing a number of significant housing differences.

In the present analysis, neighbourhoods were divided into two age categories. New areas were defined as those with 10% or less of the housing built prior to 1946¹. This essentially singles out very new housing areas. All other areas with more than 10% of the housing built prior to 1946 were lumped together as old although there is obviously a range of ages and mix of housing in this category. Descriptions of the two submarket samples are provided in Table 6. Numerous variables show marked differences between the two samples as expected. Three variables may be noted particularly. First, houses in new areas are predominantly brick or brick-face even more than in older areas. Second, virtually all houses in new areas have driveways. Third, there is virtually no industrial development in new areas. The lack of differentiation, particularly for the last two variables, means that these may not be important variables differentiating house prices in new neighbourhoods. They may, however, be extremely important in older areas.

¹These are based on the 1971 data.

To test for submarket differences the price-attribute relationship was estimated for each of the linear and log-linear models. The results are shown in Tables 6 through 10 (columns 4 and 5). Large differences are evident: space, both internal and external (rooms, recreation room, lot size) is relatively more important in new areas than old areas; internal features, as a whole are more important in older areas with some exceptions; external features are generally more important in older areas (brick, aluminum, paved driveway, garage); social characteristics (education, population, age) are very important in old areas relative to new areas; but physical characteristics are more important in new areas. The presence of multi-family housing, highrises or retailing has a significant negative impact on house prices in new areas while open space and insitutional open space have significant positive impacts on house prices. The number and size of the coefficient differences suggest that old and new areas constitute separate submarkets.

To test this more formally Chow tests were performed. The results are indicated in Table 11(b). Three of the four tests are significant at the .05 level. This provides evidence that old and new housing neighbourhoods may represent different submarkets. The fact that the log-linear model is much less decisive, however, raises some doubt. Are the results in the log-linear case a reflection of submarket differences or non-linearities in

the data? In other words, is the linear model fitting straight lines to different parts of the same curve? In any case, the results clearly demonstrate the inadequacy of a linear model using a pooled sample across different aged neighbourhoods.

The results also have an intuitive logic. The negative impact of retail facilities and highrises as well as the higher profile for open space are consistent with expectations about new areas. In contrast, social characteristics, reflecting the heterogeneity of older areas, are emphasized in these neighbourhoods. It also seems reasonable that internal and external features are relatively more important in older houses many of which suffer from a lack of modern conveniences. The evidence, therefore, while not conclusive suggests that important differences exist between new and older areas over and above differences in the mix of attributes.

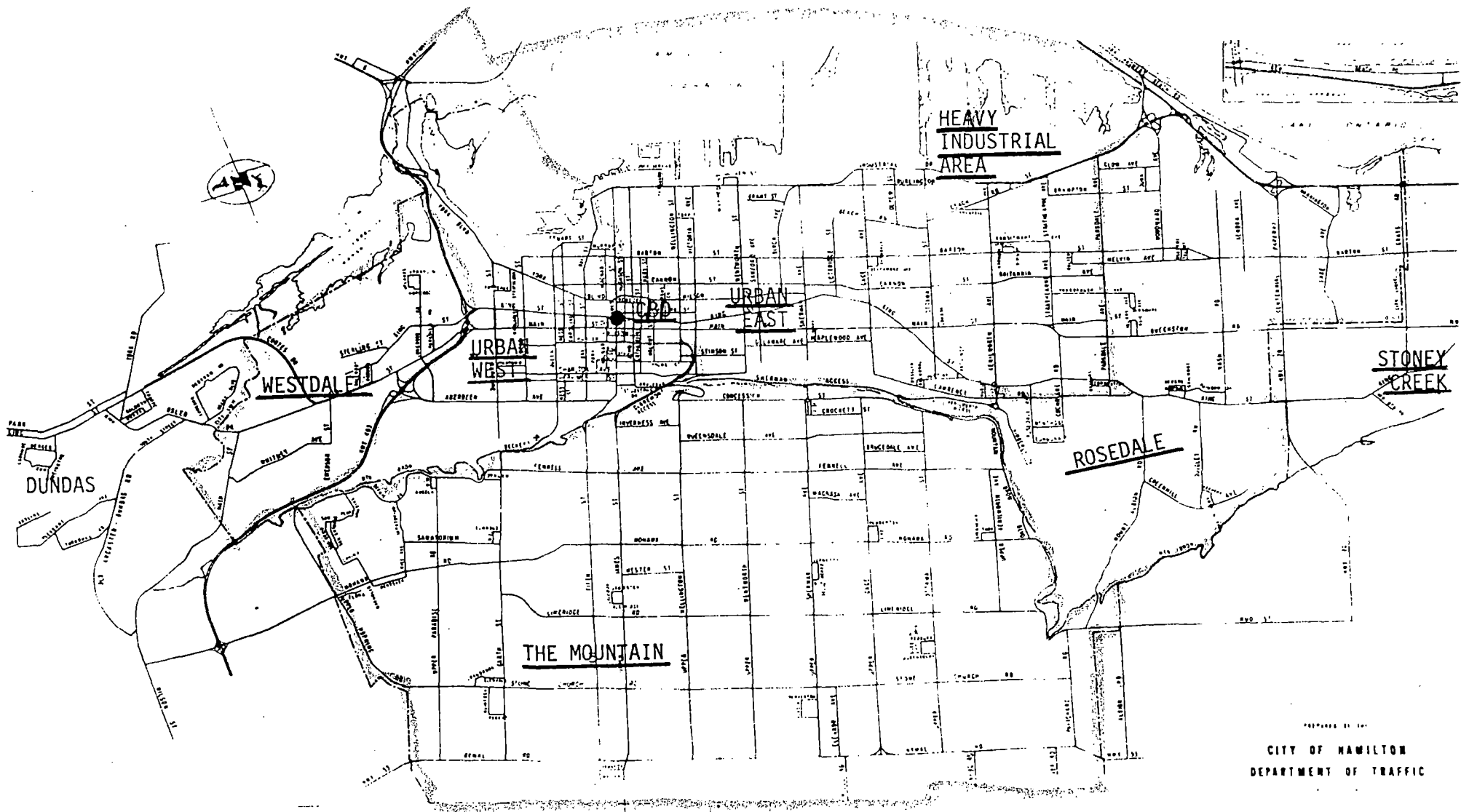
Median Income

Income is an important variable associated with differences in the social characteristics and status of neighbourhoods. To examine the potential influence of this factor in the formation of submarkets, neighbourhoods were divided into high income and low income areas. Areas with a median household income greater than \$20,000 (\$1980) were designated high income.

All other areas were low income. Descriptions of two subsamples are provided in Table 6. Differences include: lot size, most internal and external features, housing age, education and industrial and retail land use. The difference in price is very large: the average price is \$54,900 in high income areas and \$35,700 in low income areas. This reflects the importance of income as an explanatory variable.

The coefficient estimates for the low and high income submarkets are provided in Tables 7 to 10 (columns 6 and 7). Major differences are evident: internal space (rooms) is more highly valued in higher income areas; internal features such as carpeting and fireplace add more to the value of a house in a low income area although the reverse is true for air conditioning; external features, notably a paved driveway, are more valuable in low income areas but having a pool adds significant value to a house only in a high income area; a detached house is relatively more valuable in a low income area; and neighbourhood features are markedly more important in differentiating house values in low income areas, particularly the variables of industrial use, institutional open space, population age (% aged 20-34) and income.

The importance of these differences is supported by the results of the Chow tests shown in Table 11(c). All four tests are significant at the .01 level supporting the argument that high and low income areas represent separate submarkets.



PREPARED BY 101
CITY OF HAMILTON
DEPARTMENT OF TRAFFIC

Urban and Suburban Areas

Analyses of urban and suburban differences has a long tradition in urban sociology and geography. To examine differences along this dimension the area was divided into two areas. (Refer to Map 1). The area above the escarpment, known as the Mountain, was combined with the west and east suburban areas, referred to here as Westdale and Rosedale respectively to form the suburbs. These areas include some older areas, notably along the Mountain brow which distinguishes this submarket division from that of neighbourhood age. The urban area was formed by combining the areas identified as Urban East and Urban West.

Descriptions of the subsamples are provided in Table 6. Differences are evident in terms of price, lot size, most internal and external features - the suburbs being better equipped in most cases - and in several neighbourhood features notably industrial land use, age and education.

Results of the regression analyses are provided in Tables 7 to 10 (columns 8 and 9). Numerous differences are evident in comparing structural variables: internal space (rooms and recreation room) is more valuable in suburban houses; carpeting adds more to a suburban house but a fireplace is more important in an urban house; external features are generally more important in differentiating urban houses particularly the presence of a

paved driveway whereas the presence of a garage is an important characteristic in differentiating suburban houses (virtually all of which have driveways).

The most striking differences, however, refer to the neighbourhood characteristics. In suburban areas the presence of highrises and retail facilities have a significant negative impact on house prices as does age. An increase in the proportion of the population aged 20-34 has a positive impact. None of these are significant in urban areas. The key neighbourhood variables are education, institutional open space and industrial land use.

The marked differences shown are supported by the Chow test results in Table 11(d).

All tests are significant beyond the .01 level. The results provide considerable support for the conclusion that urban and suburban areas constitute distinct housing submarkets.

Regions

The rather unusual spatial structure of Hamilton particularly the sharp physical division provided by the escarpment and the major heavy industrial development along the eastern waterfront, suggests a further spatial

division for analysis. The area was divided into three regions. The Mountain was defined as one region and the area "below the escarpment" was divided into two regions. The areas of Rosedale and Urban East were combined to form the region Northeast (Map 1). The latter includes virtually all of the heavy industry in Hamilton. The CBD is on the boundary of these two regions.

Descriptions of the 3 regions are provided in Table 6. Differences are evident on a number of dimensions. The Mountain, in particular, is different from the other 2 regions. The Mountain has the highest average selling price, the highest proportion of houses with recreation rooms, carpet, brick siding and driveways; it is the youngest area with the highest level of ownership, highest income, largest lots and lowest proportion of industrial land use. In contrast, the Northeast region has the lowest average selling price, the lowest proportion with recreation rooms, carpet, brick siding and driveways; it has the lowest median income, lowest education level, and highest proportions of industrial and retail land use. The Southwest region is generally in between the other two areas except for the variables rooms, fireplaces, ownership and open space. In general, the three regions exhibit distinctly different characteristics in both structural and neighbourhood attributes.

Results of the regression analyses for these regions are presented in Tables 7 to 10 (columns 10, 11 and 12). Differences are evident on all

major groups of variables - internal and external space, internal and external features and social and physical neighbourhood attributes. Internal space (rooms and recreation room) is more valuable on the Mountain. External space, notably lot size is particularly important in the Southwest and least important on the Mountain. Internal amenities, notably carpeting and fireplace add more to the value of a house in the Northeast than elsewhere (see Table 8). A brick or aluminum exterior along with a garage add to the value of a house on the Mountain while a brick exterior along with a paved driveway add more to the value of houses in the Southwest and Northeast than on the Mountain. Neighbourhood characteristics are important in all three regions but the characteristics which are most significant differ. In the Southwest, median income is not an important variable differentiating house prices but levels of education, population age structure and housing age are important. (Table 8). As expected industrial land use has a negative impact on property values particularly in the Northeast and Southwest but not on the Mountain where industrial use is minimal. On the other hand, retail land use has a significant negative impact on house prices only on the Mountain.

To test for the significance of these differences Chow tests were performed. The results are shown in Table 11(e). In all four cases the calculated F values are significantly different from zero at the .01 level. Despite the large number of parameters estimated in this submarket model the level of explanation has been significantly improved. This is

particularly the case for the Mountain and the Southwest regions which show reductions in their standard errors. This is not the case for the Northeast. Overall, the results provide considerable support for the hypothesis that these regions represent different submarkets within the Hamilton housing market.

House Price

The final market division analyzed used the dependent variable, selling price, to define potential submarkets. Houses with selling prices less than or equal to \$40,000 were defined as Low Priced; houses with selling prices greater than \$40,000 were defined as high priced. This division is essentially arbitrary tending to isolate houses in the lower price ranges. (The average price for the sample is \$43,646). A 3 part division was examined but with a small sample of houses in the upper price ranges (over \$60,000) comparisons were limited. The basic question is: does the lower end of the housing market behave differently from the middle and upper end?

Comparison of the characteristics of the two subsamples are provided in Table 6. The price difference - \$27,900 versus \$54,900 - is very large indicating the relatively large number of very low prices houses in the sample. Comparing the structural variables, every variable (except

proportion with aluminum siding) is lower for the low price sample. This is particularly noticeable for lot size and all the internal and external features. Exceptions are the variables rooms, aluminum and detached. Differences among the neighbourhood variables include income, education, housing age, industrial land use and retail land use. In all cases the low priced sample is less well endowed. There are, therefore, large differences in the attribute mix of these subsamples as one would expect. But the question is: are the attributes valued differently in the two submarkets.

The regression results are provided in Tables 7 to 10 (columns 13 and 14). Major differences are evident. With respect to internal and external space neither "number of rooms" nor "lot size" are significant in the low price sample. This is in sharp contrast to the high price submarket, and virtually all other submarket analyses. This means that internal and external space are not significant variables in differentiating house values in the lowest segment of the housing market. Expressed differently, a buyer limited to buying a house in the low end of the market does not appear to use number of rooms or lot size as important criteria in choosing between houses. This is surprising but consistent with the concept of housing submarkets.

Other differences are evident, carpeting is particularly valuable for low priced but not high prices houses, while a recreation room, fireplace and

air conditioning add significantly to the value of high priced but not low priced houses. For low priced houses virtually all external features are important while for high priced houses only brick siding and a garage add significantly to their value. In general, the value of internal and external features differs significantly between these market segments.

Differences in neighbourhood variables are also marked. Six of the neighbourhood variables in the "best fit" log-linear model (Table 8) and seven in the linear model (Table 7) are significant for the high priced sample but only three and four variables respectively in the low priced market. Particularly important in the high priced submarket are variables relating to social characteristics and status. As well as income and population age structure, housing age, multi-family units and highrises also have social implications.

To test for the significance of these differences, Chow tests were performed. The results are presented in Table 11(f). All tests are significant at the .01 level. The improvement of this partition is reflected in the lower standard errors and high overall level of explanation reflected in the combined R^2 statistic. The individual submarket R^2 values are low reflecting the reduced variation in the variables using this partition. Altogether, the results provide strong support for the hypothesis that lower priced houses represent a distinctly different submarket from medium and high priced houses.

VI. SUMMARY AND CONCLUSION

The purpose of the study has been to analyze the hypothesis that urban housing markets are not single well behaved markets but segmented into a number of quasi-independent submarkets. Submarkets are thought to arise because both the supply and demand for housing is highly segmented. This means that there is little substitution between house types: buyers with particular demands tend to focus on one segment of the market without comparing across market segments. This creates the situation in which each housing segment - intersection of supply and demand - forms a distinct submarket only loosely connected to other submarkets. In other words, at any moment in time different market segments may be experiencing different supply-demand conditions.

The analytical effect of varying submarket conditions is that the relationship between the selling price of a house and its set of attributes is affected. The implicit value of an attribute - on additional rooms, the presence of a fireplace or garage, or a neighbourhood free of industrial or retail activity - may vary from one submarket to the next. Submarkets may exist without such differences at any moment in time, but if such differences do exist then this provides evidence that submarkets may be operating.

Given this logic, the analysis has involved three steps. First, a series of potential submarket divisions were defined. Second, for each submarket division the price-attribute relation was estimated for each submarket. Finally, the coefficient estimates which reflect the implicit value attached to attributes within each submarket were compared across submarkets.

Six submarket divisions were analyzed. Submarkets were defined on the basis of house size (number of rooms), neighbourhood housing age, neighbourhood median income, urban and suburban areas, spatial regions, and market selling price. Only for the division based on housing size was the submarket hypothesis conclusively rejected. Strongest evidence supporting submarkets was provided for urban/suburban submarkets, regional submarkets and selling price submarkets.

There was considerable overlap in the nature of the submarket differences across submarket partitions. Older urban areas with lower incomes and lower selling prices typically placed much greater emphasis on internal and external features, e.g. carpeting, fireplace, and paved driveway and social characteristics of the neighbourhood, e.g. income, education and population age. The ability to take over an existing mortgage as well as the negative impact of industrial land use in an area were also important. In contrast, newer suburban areas with higher incomes and higher selling prices placed

much greater emphasis on internal and external space, i.e. number of rooms and lot size, and physical neighbourhood characteristics notably the extent of institutional open space, multi-family and high rise accommodations and retailing. Specific internal and external features were also more important in these areas, particularly the presence of air conditioning and garages. The presence of a paved driveway was notably unimportant in differentiating the value of houses in this area because virtually all houses had this feature. Industrial use was also unimportant for the same reason. Variables which showed little variation across submarkets were the number of bathrooms and the presence of a brick exterior. The latter, in particular, was an important variable influencing the selling price.

Overall, four broad conclusions may be drawn from the analysis. First, there is considerable evidence that distinct market segments exist in an urban housing market consistent with the hypothesis of submarkets.

The strongest evidence is provided in Tables 7 to 10 which show a tremendous variation in the magnitude of coefficients across submarkets. It is clear that pooling a sample across structural, neighbourhood and locational types and analyzing the price-attribute relationships for this pooled sample masks important market segment differences. A disaggregate approach as undertaken here provides a great deal more information about how "the" market is operating.

The second conclusion, however, is that in many instances the cost and effort of a disaggregate approach is unnecessary. If one is concerned with developing a predictive model of housing prices, perhaps as part of a larger study then the single market model performs almost as well as the disaggregate models. Although the submarket model significantly improves the level of explanation statistically, in real terms the improvement is not dramatic. (Although the division by price is approximately 10% which is large). In contrast, if the objective is to use the analysis to devise a marketing strategy for a real estate firm or estimate the market value of adding a specific feature (e.g., fireplace or garage) then a disaggregate approach is important. For example, it would seem to be important for a real estate firm to know in its advertising that a garage is an important differentiating factor for a house in a suburban area but a paved driveway is more important in an older urban area. Nevertheless, for many purposes the single market model provides an excellent account of the data.

The third conclusion is that the log-linear form of the model provides consistently better fits to the data than the linear form. This is consistent across all analyses and supports the theoretical expectation that some of the relationships are non-linear.

Finally, fourth, the importance of incorporating social neighbourhood characteristics has been demonstrated. Variables such as level of

education and population age structure are frequently significant variables influencing housing prices as well as income which has been traditionally included.

The question of the extent and importance of submarkets has not been resolved in any conclusive way in this study. The analysis does, however, support the existence of submarkets and suggests that their impacts are important, in certain types of analysis.

REFERENCES

- Ball, M.J. (1973) "Recent Work on the Determinants of Relative House Prices", Urban Studies, 10, 213-33.
- Ball, M.J. and R.M. Kirwan (1977) "Accessibility and Supply Constraints in the Urban Housing Market", Urban Studies, 14, 11-32.
- Brummell, A.C. (1981) "A Test of Spatial Submarkets in Urban Housing", Can. Journ. of Reg. Sc., IV, 89-111.
- Evans, A.W. (1973) The Economics of Residential Location. London: Macmillan Press.
- Grother, D.M. and P. Mieszkowski (1974) "The Determinants of Real Estate Values", Journ. of Urban Economics, 1, 127-46.
- Jackson, J.R. (1979) "Intraurban Variation in the Price of Housing", Journ. of Urban Economics, 6, 464-79.
- MacLennan, D. (1977) "Some Thoughts on the Nature and Purpose of House Price Studies", Urban Studies, 14, 59-71.
- Mark, J. (1977) "Determinants of Urban House Prices: A Methodological Comment", Urban Studies, 14, 359-63.
- Mills, E. (1969) "The Value of Urban Land", in H. Perloff (ed.), The Quality of the Urban Environment. Washington, D.C.: Resources for the Future.
- Nelson, J.P. (1975) The Effects of Mobile-Source Air and Noise Pollution on Residential Property Values. Report for the Dept. of Transport, Washington, D.C.
- Richardson, H.W., J. Vipond and R. Furbey (1974) "Determinants of Urban House Prices", Urban Studies 11, 189-99.
- Rosen, S. (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition", Journal of Political Economy 3, 146-66.
- Schnare, A.B. and R.J. Struyck (1976) "Segmentation in Urban Housing Markets", Journal of Urban Economics, 3, 146-66.
- Silk, J. (1979) Analysis of Coyariance and Comparison of Regression Lines. Concepts and Techniques in Modern Geography, 20. Norwich, England: Geo Abstracts Ltd.

Smit, B. (1979) "Covariance Methods for Spacio-Temporal Analysis",
Professional Geographer, XXX, 140-48.

Straszheim, M.R. (1975) An Econometric Analysis of the Urban Housing Market. Urban and Regional Studies No. 2 New York: Bureau of Economic Research.

Zarembka, P. (1974) "Transformation of Variables in Econometrics", in P. Zarembka (ed.). Frontiers of Econometrics, New York: Academic Press.