Volume Title: An Econometric Analysis of the Urban Housing Market

Volume Author/Editor: Mahlon R. Straszheim

Volume Publisher: NBER

Volume ISBN: 0-870-14512-6

Volume URL: http://www.nber.org/books/stra75-1

Publication Date: 1975

Chapter Title: A Model of Housing Prices and Neighborhood Incomes

Chapter Author: Mahlon R. Straszheim

Chapter URL: http://www.nber.org/chapters/c0982

Chapter pages in book: (p. 142 - 165)

6

A MODEL OF HOUSING PRICES AND NEIGHBORHOOD INCOMES

1. INTRODUCTION

The sharp discontinuities in any geographic mapping of neighborhood incomes and prices arise from the discontinuities over space in demand and supply functions for housing. Jurisdictional boundaries affecting the level of public services and taxes, racial neighborhood boundaries, differences in accessibility to shopping and employment, and the heterogeneity in the existing housing stock all affect demand and supply functions. In the model below, these factors characterizing particular housing submarkets are regarded as exogenous at a point in time. Using an econometric approach, the model analyzes the market for fixed residential stocks. The disaggregation of the housing market is that used in Chapter 3; the Bay Area is divided into eighty-one geographic areas, each including several submarkets for particular types of units. It is hardly surprising that neighborhood income is correlated to housing prices; the difficult task is to explain why high-income households outbid low-income households for a particular type of housing in any given area. The model addresses this interdependency.

2. MODEL SPECIFICATION

Characteristics of the capital stock used in defining submarkets include the percentage of single-family structures, structure age for single- and multifamily units, and lot size for single-family structures. Age serves as a proxy for a variety of characteristics of the stock, such as room size and layout. Each of these attributes have been shown to be important in households' preferences for housing.

The model assumes that employment access, housing-stock availability, and neighborhood racial composition are exogenous; income and housing prices are considered endogenous.

Exogenous:

D = matrix describing neighborhood desirability (d_{ji} = attribute i in neighborhood j):
- LT = average lot size (acres);
- AG = average age of owner-occupied units (years);
- AGR = average age of renter-occupied units (years);
- EA = employment access (see below).

R = matrix denoting racial composition (r_{ji} = racial composition in neighborhood j measured by dummy variable i):
- R_1 = race dummy (1 if area 15 to 60 percent black, 0 otherwise);
- R_2 = race dummy (1 if area more than 60 percent black, 0 otherwise).

S = matrix describing stock availability (s_{jk} = amount of house type k available in neighborhood j).

Endogenous:

Y = matrix of occupant incomes (y_{jk} = income of occupants of house type k in neighborhood j).

P = matrix of housing prices (p_{jk} = price [or rent] of house type k in neighborhood j).

The definition of several of the independent variables deserves comment. Employment access at zone i is a weighted sum of travel times to all jobs, the weights determined by employment concentration at each site j.

$$EA_i = \sum_{j=1}^{n} \left[V_j T_{ij}^{-b+1} \Big/ \sum_{j=1}^{n} V_j T_{ij}^{-b} \right],$$

where

V_j = level of employment at site j; and
T_{ij} = travel time between i and j.

Employment access is measured in terms of average travel times, with a low value denoting high access.[1] This gravity model formulation is a substantially better way to represent spatial variation in activity than simply using "distance to the center." The race dummies denote virtually all-black areas or mixed-racial areas; the latter may or may not be in the process of tipping to all-black occupancy.[2] Two types of variables describe the supply of units available: the percentage of the total stock which consists of single-family structures and the percentage of the stock of single- or multifamily units represented by each age or lot-size class. Since the zones have approximately the same number of housing units, these two variables fully describe the supply side.

The above assumptions regarding the specification of exogenous variables deserve justification. The employment access variable (*EA*) reflects job locations and the transport system. While population densities influence certain employment location decisions, occupant incomes and prices for a *given* distribution of housing units seem unlikely to have significant causal effects on employment locations or travel times to employment centers at any given point in time.[3] The vagaries in the planning of highways and public transit systems are notorious. Inattention or the inability to predict markets, long time lags in decision making and implementation, and the inability to alter capital infrastructure once in place all serve to insure that the performance of the transport system

1. This form of accessibility variable is familiar in land use–transportation studies. Average cost is assumed to be a weighted sum of the travel times to all other zones, weighted by the probability a trip is made. The latter is represented by a gravity model:

$$\left[V_j T_{ij}^{-b} \Big/ \sum_{j=1}^{n} V_j T_{ij}^{-b} \right]$$

Employment locations and travel times and the accessibility variable were based on a survey taken by the Bay Area Transportation Commission.

2. Definition of the cutoff points of 15 percent and 60 percent was based on minimizing the variance of the variable "percent black occupancy" within the three classes defined by the race dummies. No neighborhoods were close to the cutoff points.

3. Causation from income to employment levels is negligible. It does not require appreciably more employment to service high-income households, despite the fact that they spend more. Also, evidence indicates that higher-income households travel farther to shop, aided by higher levels of automobile ownership. In any given cross section, there is a substantial variance in these jobs' locations as a function of residential locations because of the costs of firm relocation. Retailing and service sectors account for less than half of total employment.

can be regarded as exogenous in a cross-section model of the sort outlined here.

The assumption that the spatial pattern of racial discrimination is exogenous is based on the stability over time of the racial composition of many neighborhoods (except for that small subset of markets which are tipping). The changes in segregation barriers which occur in response to changes in housing-market conditions, changes in incomes and in prices of housing at the boundary, the growth of demand in the two markets, and so on, are highly complex and indirect.

The third assumption of a fixed supply of housing by type in any given location at a point in time is based on the very long lags in housing-stock adaptations. Some changes in demand for housing may be anticipated and new supply provided or existing housing modified or destroyed. Supply responses are most rapid where vacant land is available; yet, in the fastest-growing suburbs in San Francisco only 20 percent of the stock is less than five years old. Rapidly growing suburbs often exhibit sharply rising prices because of lags in supply adjustments. Falling prices and high vacancy rates are a familiar sight in some neighborhoods in the central portions of many cities where the demand for housing has sharply declined.

Two basic types of equations are employed to explain income and price in a given submarket. The first relates housing prices of house-type k in neighborhood j to the income of occupants, the desirability and racial composition of the area, and the supply of units. It is derived from the market's demand curve and an expression for market equilibrium. Higher income of occupants or a neighborhood more accessible to employment should increase people's bids, shifting up the demand curve. The race dummies in the price equation measure the differential that residents of black neighborhoods will pay, independent of income and employment accessibility. These coefficients will reflect any differences in tastes between ghetto and nonghetto residents, the premium or discount which the market places on residing in a black submarket by reason of its racial composition, and any differences in the bidding process for housing between the white and black submarkets. A neighborhood with a higher percentage of whites may be preferred by many households. If those living in the white submarket strongly prefer a segregated market, the price in white submarkets will be bid up (discouraging entry by those indifferent to the racial mix). This implies a negative race dummy. Second, price competition may be different in the black submarket than in the white submarket. If housing market discrimination restricts black households choices but not those of whites, the supply variables in the price equation may not equally represent the supply available to households bidding in the two markets. For example, competition may bid up the price for the types of units in short supply in the black submarket. Conversely, for the types of units in overabundant supply in black neighborhoods relative to white areas of comparable income, the race coefficient may be negative. A third possibility is that the race dummies reflect neighborhood characteristics which have been omitted from the equation.

$$\text{Demand equation } q^D_{jk} = f(P,D,R,y_{jk}). \tag{1a}$$

$$\text{Market equilibrium } q^D_{jk} = s_{jk}. \tag{1b}$$

Since there are no vacancies in the model, the quantity demanded in each submarket q^D_{jk} must just equal the supply s_{jk}. Solving for p_{jk} yields the equation for prices:

$$P_{jk} = g(EA,R,P_{j'k'}, y_{ik}, s_{jk}), \tag{1}$$

where $P_{j'k'}$ denotes the vector of prices for all other submarkets except jk. Other submarkets may be substitutes, and they may produce external effects on prices as well.

Since, in this model, the determinants of the demand curve are assumed to be independent of the supply of housing in each submarket, prices must change if each submarket is to clear. The assumption of a fixed stock in the short run which is not fully adapted to current or anticipated demand at any given point in time implies that observed prices for particular types of housing may bear little relation to their long-run supply prices. Across submarkets, "temporary" demand and supply variations produce substantial variations in quasi rents to those supplying the housing stock. This can lead to either abnormal profits or losses. Changes in the utilization of the capital stock were omitted because no data on vacancy rates were available.

The second equation relates occupant incomes of house-type k in neighborhood j to prices, neighborhood desirability, and racial composition, reflecting, in essence, how the existing stock of housing was allocated among households:

$$y_{jk} = f(D,R,P). \tag{2}$$

The signs of $\partial y_{jk}/\partial d_{jik}$ will reflect household tastes and incomes. Income elasticities for age and lot size are significant, and hence, lot size and age variables should have positive coefficients as high-income households outbid lower-income households. Access to employment should be more highly valued by high- than by low-income residents, assuming that the value of travel time is positively related to income. There is ample documentation indicating that higher-income families commute farther but acquire more housing. Finally, prices should affect who chooses what neighborhood, with higher prices likely to rule out certain groups, in this case, assumed to be those with lower incomes.

The race coefficients in the income equations measure two effects. Black household incomes are below those of white households. Moreover, in addition, racial discrimination alters the residential choices of both blacks and whites. Those black households which obtain access to the white submarket are predominantly middle- and upper-income blacks. At the same time, many whites alter their location and housing choices in order to escape integrated or predominantly black neighborhoods. Average incomes of white households residing in the black submarket is very low. Each of these factors implies a negative coefficient for the race dummies in the income equation.

The form in which independent variables denoting income, neighborhood desirability, and prices for other submarkets than the one in question enter Equations 1 and 2 is more complex. Much depends on the particular geography of the city; i.e., whether areas are sufficiently near each other to be viable substitutes or are such that prices or incomes in one exert a neighborhood effect on an adjacent area. For example, the existence of high-quality neighborhoods may raise incomes and prices in adjacent areas. At any given point in time, there exists a fixed set of housing stocks, a given distribution of income and consumer tastes, and fixed employment locations. The analytic form of functions 1 and 2 will therefore reflect the distribution of income and tastes, the availability of units by type and neighborhood, and the nature of the bidding process.

This level of generality virtually precludes an analytic solution for an equilibrium in all submarkets. As has been noted earlier, existing monocentric models consider only one market, the market for land, and employ highly simplifying assumptions in order to obtain closed-form solutions. If space is treated in rectangular coordinates rather than by continuous gradients and if neighborhood boundaries are discrete, the familiar marginal conditions for household equilibrium often lead to corner solutions. Obtaining closed-form solutions becomes very difficult. Introduction of many types of housing further complicates the specification. Programing procedures of a more ad hoc sort are typically employed in urban land-use models to clear markets[4] though if there are many types of housing and locations to choose from, specifying the objective

4. James Brown et al., *Empirical Models of Urban Land Use* (New York: NBER, 1972).

function in any optimization routine also generally involves many arbitrary assumptions.

In simplifying the specification to permit econometric estimation, the spatial interdependencies among submarkets must be neglected. The model estimated below uses a linear form for Equations 1 and 2 for each house-type k, and is estimated using data across geographic submarkets j. Each submarket j is assumed to provide one observation indicating how prices and incomes change when values for the exogenous variables differ. No substitute or complementary terms are included in the estimation below; independent variables include only income, racial composition, and neighborhood characteristics of the submarket in question. These linear empirical expressions are a simple representation of the effects of changes in exogenous variables on incomes and prices in the sample region. The monocentric models treat spatial interdependence in a single market (for land). My approach analyzes market conditions across a large number of submarkets for different types of housing which are recognized to vary in an irregular fashion over space. However, this analysis is achieved by forgoing an explicit treatment of the spatial interdependencies in these submarkets.

The model includes an equation for income and one for prices for each of ten particular structure types, as well as equations for average income and average prices for all rental units and all owner-occupied units. (The twenty-four equations are shown in Table 6.2).

The treatment of both prices and incomes as endogenous is one of the principal characteristics distinguishing this model from other regression studies of housing prices. The structure of the model allowing identification of the parameters of both the income and price equations can be briefly summarized. The price equation includes as exogenous variables, supply, race, and employment access. Other characteristics affecting a neighborhood's desirability are excluded, as they are represented by the income variable. If it can be explained why high-income households live in one area and not another, prices can be well predicted. In explaining incomes by neighborhood and stock type, all of the neighborhood characteristics are relevant, along with prices; the supply of units (independent of prices) is irrelevant to households' choices.

Overidentification is achieved by omitting supply variables from the income equation and omitting most neighborhood variables from the price equations. This is apparent in the form of the equations for particular types of units. The price equation typically includes race, supply variables, as denoted by the percent of single-family structures and the percent of units of the particular age or lot size in question, and employment access. The income equation includes employment access, race, and additional neighborhood desirability variables, but excludes any supply variables. The exception to this format is the case of the equations for prices and incomes aggregated *across* all housing types. To account for this aggregation, average age and lot size, two neighborhood desirability variables, were included in the price equation as well as in the income equation. Overidentification of the two equations results from the vector of supply variables of particular types of units being excluded from each equation, as well as the percent of single-family-units variable being excluded from the income equation.

While the model is principally concerned with explaining neighborhood incomes, three alternative models were estimated employing different assumptions regarding the role of education. The three alternative specifications of the model are illustrated by reference to two of the equations in the entire set, the equations for average income and average prices for owner-occupied units.

Model 1

The first model is as described above, excluding education:

$$YT = f(EA, LT, AG, PT, R1, R2); \tag{3}$$

$$PT = g(YT, EA, OWN, R1, R2). \tag{4}$$

Model 2

The major omission from Model 1 is the quality of public services and tax burdens, which are also likely to be important in households' location choices. The one characteristic of activity in the local public sector for which data is available is sixth-grade reading scores by zone. Unfortunately, the lack of any data on local education inputs raises serious problems in determining the appropriate specification. In brief, it is not possible to assess whether education test scores reflect parents' education, the quality of the local education process, or both.

One approach would introduce education scores as endogenous. Including education test scores in the income equation might reflect the point that test scores are one measure of neighborhood desirability, or it might reflect the fact that children from higher-income households generally score better on achievement tests. Equation 3 now becomes:

$$YT = f(ED, EA, LT, AG, PT, R1, R2). \tag{3a}$$

In specifying an equation to explain education test scores, two influences should be included: the quality and amount of public education inputs, and the preschool, family, and other environmental circumstances which affect students' ability to learn. Lacking data on local education inputs, the simplest version would specify only income and race as explanatory variables.

$$ED = f(Y, R1, R2). \tag{5}$$

As will be seen below, this specification incorporating the interdependency between neighborhood incomes and test scores has little effect on the reduced-form results.

Model 3

The current human-capital literature suggests that the appropriate specification would introduce parental education as an explanatory variable in Equation 5. Educators generally stress parents' education as an important factor influencing student achievement, given any level of education inputs.[5] Parents' education has proven to be a good predictor of first-grade IQ or achievement scores, which in turn prove relatively highly correlated with reading or achievement scores in later grades. Parents' education would thus be included as the explanatory variable.

$$ED = f(EDP, R1, R2). \tag{5a}$$

Income, as well, may exert a separate effect on education. Because income and parental education are correlated (.512), a conclusive estimate of the effects of each cannot be made. However, based on other studies, the a priori argument for including parental education is convincing. In addition, experimentation with the equation revealed that the parental-education variable in this sample is more than a proxy for income. Equation 5a, including parental education alone, exhibited a significantly higher R^2 than Equation 5, using income alone. When both factors were included, the income variable proved insignificant and the coefficient for parental education was only modestly altered.[6]

5. Christopher Jencks, *Inequality: A Reassessment of the Effect of Family and Schooling in America* (New York: Basic Books, 1972).
6. The estimates were:

$$ED = \underset{(4.07)}{23.69} + \underset{(5.04)}{.00243\,YT}; \qquad R^2 = .2434$$

$$ED = \underset{(6.01)}{36.37} - \underset{(3.26)}{.00157\,YT} - \underset{(3.60)}{17.00R1} - \underset{(2.95)}{18.83\,R2}; \qquad R^2 = .3911$$

$$ED = \underset{(1.70)}{23.61} + \underset{(4.70)}{7.713\,EDP} - \underset{(.92)}{.0009\,YT} - \underset{(2.51)}{12.01\,R1} - \underset{(2.60)}{16.33\,R2}. \qquad R^2 = .5401$$

The parental-education variable in this equation may well reflect other factors besides students' ability to achieve, namely, differences in local education inputs. Parents with more education are likely to choose areas with good school systems, or areas where they feel they can influence the education process. Hence, while Models 2 and 3 treat education test scores and incomes as endogenous, there is insufficient data to test the separate causal factors affecting the level of test scores. In Model 3, the questions remain: (1) Why do educated parents tend to locate in particular areas? (which the model treats as exogenous); and, (2) Does parental education represent students' ability to learn or is it a proxy for something concerning the education process in local communities?

3. ESTIMATION OF THE MODEL

A. *The Data*

The model requires data on occupant income and housing prices by type of housing and by neighborhood. Occupant incomes were reported in the household-interview survey. Prices for different types of housing were based on the estimated hedonic price indexes for each submarket described in Chapter 3.

B. *Estimation of the Structural Equations*

The models are estimated by two-stage least squares, using a sample of eighty-one zones. For purposes of comparison, estimates of the three equations for average income, average prices, and education for the three models are shown in Table 6.1. The quality of fits are about comparable. There is little difference in the reduced forms for Models 1 and 2 which, respectively, exclude and include education test scores as endogenous. The introduction of parental education in Model 3 alters the results, primarily by reducing the coefficients on the dummies for neighborhood racial composition. This is the expected result, given the differences in parents' education by race. The significance of the race coefficients in the education equation could reflect neighborhood differences, a pure race effect, or differences in education inputs. This last may be important if black communities receive less than their share of education inputs. The reduced-form coefficients for the other variables are very similar for all three models.

Tables 6.2 and 6.3 show the complete set of structural equations for each housing type for Model 3. The owner equations exhibit the better statistical fit, with most of the coefficient exhibiting the expected sign and many of the *t*-ratios approaching 2 or higher. In the income

TABLE 6.1
STRUCTURAL AND REDUCED-FORM EQUATIONS FOR AVERAGE OWNER INCOMES AND PRICES: MODELS 1, 2, AND 3

A. Structural Equations

Model 1.	*YT* =	7631	+.1541 *PT*	−45.33 *EA*	+9495 *LT*	−111.9 *AG*	−2553 *R1*	−3261 *R2*		R^2 = .6332
		(7.09)	(3.04)	(1.40)	(3.98)	(1.28)	(2.92)	(2.65)		
	PT =	20050	+.8501 *YT*	−150.1 *OWN*	−95.07 *EA*	+3005 *LT*	−204.1 *AG*	−5200 *R1*	−8505 *R2*	R^2 = .6821
		(1.39)	(2.05)	(1.98)	(.99)	(2.04)	(1.88)	(.85)	(1.72)	
Model 2.	*YT* =	−3430	+.2252 *PT*	+190.5 *ED*	−30.40 *EA*	+6024 *LT*	−196.6 *AG*	+1418 *R1*	+937 *R2*	R^2 = .5149
		(.78)	(3.49)	(2.65)	(.80)	(1.97)	(1.85)	(.78)	(.94)	
	PT =	20050	+.8501 *YT*	−150.1 *OWN*	−95.07 *EA*	+3005 *LT*	−204.1 *AG*	−5200 *R1*	−8503 *R2*	R^2 = .6821
		(1.39)	(2.05)	(1.98)	(.99)	(2.04)	(1.88)	(.85)	(1.72)	
	ED =	45.48	+.0008 *YT*	−17.92 *R1*	−17.64 *R2*					R^2 = .3640
		(4.25)	(.91)	(3.52)	(2.66)					
Model 3.	*YT* =	1131	+.1746 *PT*	+116.5 *ED*	−40.22 *EA*	+7339 *LT*	−128.3 *AG*	−311 *R1*	−971 *R2*	R^2 = .7111
		(.64)	(4.56)	(3.85)	(1.41)	(3.32)	(1.85)	(.85)	(.85)	
	PT =	24801	+1.105 *PT*	−396.0 *OWN*	−56.01 *EA*	+4002 *LT*	−208.0 *AG*	−2205 *R1*	−350.4 *R2*	R^2 = .7405
		(2.45)	(3.44)	(2.11)	(.91)	(2.15)	(2.23)	(1.58)	(2.05)	
	ED =	22.27	+6.656 *EDP*	−10.71 *R1*	−14.54 *R2*					R^2 = .5282
		(1.71)	(6.00)	(2.48)	(2.58)					

B. Reduced-Form Results

		Exogenous						
	Endogenous	*EA*	*PT*	*AG*	*OWN*	*R1*	*R2*	*EDP*
Model 1.	*YT*	−68.5	11450	−164.9	−16.5	−3856	−5224	[a]
	PT	−153.6	12755	−344.4	−172.5	−8479	−12970	[a]
Model 2.	*YT*	−72.3	9044	−333.7	−50.0	−4396	−6097	[a]
	PT	−153.0	10051	−464.0	−192.0	−11260	−15847	[a]
	ED	−.054	1.35	−.07	−.037	−21.65	−20.63	[a]
Model 3.	*YT*	−59.4	9818	−201.8	−85.2	−2416	−2810	959
	PT	−110.3	14797	−421.9	−489.9	−4468	−7965	1060
	ED					−10.8	−14.5	6.66

Note: Numbers in parentheses are *t*-ratios.

[a] Not applicable to model.

equations, prices and education coefficients have the highest *t*-ratios, followed by lot-size and age variables. Employment access and race coefficients have the expected signs but have lower *t*-ratios. These equations are a significant improvement over the simple empirical gradient functions presented in Chapter 3. In the price equations, income, race, and the tenure mix of the neighborhood generally prove significant. Employment-access coefficients assume the correct sign but have low levels of significance. The variable denoting the supply of housing of a given type has the correct sign only in the case of the two equations for units built since 1950.

Elasticities (evaluated at the mean) of neighborhood incomes are highest with respect to prices (about .75, depending on the structure type), next highest with respect to education (about .4), and lower with respect to lot size, age, and employment access. The elasticity of neighborhood income with respect to prices varies by stock type and is lower for the newest units and for those with larger lots. The fact that the rationing effect of prices on incomes of those occupying the newest units is lower may reflect a situation in which higher-income households are bidding for their occupancy.

The rental submarket equations are of the same form. In the rental income equations, prices, education, structure age, employment access, and race all prove significant. Elasticities with respect to prices are about .50, and are about .30 with respect to education, structure age, and employment access. Race exhibits a positive effect on renter incomes, holding rents and neighborhood quality constant. This probably reflects the fact that many middle-income black households are forced to rent rather than own because of the shortage of single-family structures in the ghetto. The equations for rents by structure type are less reliable. Only income has a high *t*-ratio; the coefficients imply elasticities of about .80. The variable for percentage of one-family structures has a negative coefficient, implying that a higher percentage of such structures raises rents. This variable most likely serves as a proxy for the tenure mix or neighborhood stability rather than as a supply-curve variable. The employment-access variable is insignificant. Only a few of the race coefficients have *t*-ratios exceeding one; these estimates imply that race exerts a negative effect on rents if income and employment accessibility are held constant.

C. *The Reduced-Form Results for Model 3*

Solution of the equations for incomes and prices in terms of the vector of exogenous variables reveals the total effect of a change in any exogenous variable. Coefficients of the reduced-form equations are shown in Table 6.4, together with elasticities evaluated at the sample

TABLE 6.2
STRUCTURAL EQUATIONS: OWNER OCCUPANTS

Equation Number										
1	YT =	+1131	+.1746 PT	+116.5 ED	−40.22 EA	+7339 LT	−128.3 AG	−311 R1	−971 R2	R^2 = .7111
		(.64)	(4.56)	(3.85)	(1.41)	(3.32)	(1.85)	(.85)	(.86)	
2	Y3 =	−4023	+.3253 P3	+102.4 ED	−31.38 EA	+7454 LT		+150 R1	+275 R2	R^2 = .5375
		(1.37)	(5.24)	(2.34)	(.73)	(2.23)		(.11)	(.16)	
3	Y4 =	−2318	+.3762 P4	+45.57 ED	−33.05 EA	+4125 LT		−91 R1	−632 R2	R^2 = .4523
		(.85)	(6.37)	(1.02)	(.72)	(1.23)		(.07)	(.89)	
4	Y5 =	−4837	+.4149 P5	+112.3 ED	−46.52 EA	+5771 LT		−84 R1	−1747 R2	R^2 = .4940
		(1.51)	(7.16)	(2.11)	(.88)	(1.43)		(.05)	(.91)	
5	Y6 =	−6522	+.3357 P6	+199.2 ED	−15.53 EA	46 LT		−515 R1	−2378 R2	R^2 = .4678
		(1.35)	(6.16)	(2.42)	(.99)	(.01)		(.21)	(.79)	
6	YS =	4239	+.1717 PS	+78.14 ED	−30.82 EA		−124.0 AG	−70 RI	−896 R2	R^2 = .5424
		(2.52)	(4.17)	(2.68)	(1.39)		(2.03)	(.08)	(.83)	
7	YL =	2011	+.3239 PL	+62.78 ED	−40.81 EA		−87.2 AG	−1176 R1	−5387	R^2 = .4615
		(.39)	(2.56)	(.70)	(.60)		(.46)	(.46)	(1.63)	
13	PT =	24,801	+1.105 YT	−396.0 OWN	−56.01 EA	+4002 LT	−208 AG	−2205 R1	−3504 R2	R^2 = .7405
		(2.45)	(3.44)	(2.11)	(.91)	(2.15)	(2.23)	(1.58)	(2.05)	
14	P6 =	37,250	+1.076 Y6	−213.4 OWN	−10.30 EA	−365.8 S6		−34 R1	−384 R2	R^2 = .6516
		(3.08)	(2.35)	(2.62)	(.08)	(2.50)		(.81)	(1.05)	
15	P5 =	23,150	+1.22 Y5	−37.0 OWN	−81.74 EA	−176.5 S5		−2023 R1	−3150 R2	R^2 = .5533
		(2.65)	(2.85)	(.64)	(.96)	(2.54)		(1.58)	(1.60)	
16	P4 =	22,370	+.901 Y4	−72.71 OWN	−192.1 EA	+122.2 S4		−253 R1	−2967 R2	R^2 = .5686
		(2.05)	(3.48)	(1.13)	(2.51)	(1.20)		(.08)	(1.97)	
17	P3 =	20,380	+925 Y3	−100.3 OWN	−35.33 EA	+69.8 S3		−3617 R1	−7543 R2	R^2 = .5372
		(3.11)	(2.12)	(1.93)	(.47)	(1.85)		(1.18)	(1.86)	
18	PS =	8,105	+2.928 YS	−278.9 OWN	−64.09 EA	+91.93 SS		+2258 R1	+2321 R2	R^2 = .5747
		(1.84)	(3.34)	(4.19)	(.50)	(1.27)		(1.54)	(1.38)	
19	PL =	28,240	+1.301 YL	−359.0 OWN	−82.10 EA	−116.0 SL		+5707 R1	+8350 R2	R^2 = .6957
		(1.84)	(4.70)	(3.02)	(.40)	(2.77)		(1.05)	(2.06)	
25	ED =	22.27	6.656 EDP	−10.79 R1	−14.54 R2					R^2 = .5282
		(1.71)	(6.00)	(2.48)	(2.58)					

Notes to Tables 6.2 and 6.3:
Figures in parentheses are *t*-ratios.
Y = matrix of occupant incomes

YT = average income of all owner-occupants
Y3 = income of occupants of houses built before 1939
Y4 = income of occupants of houses built before 1940–49
Y5 = income of occupants of houses built before 1950–59
Y6 = income of occupants of houses built before 1960–65
YS = income of occupants of houses built on lot .2 acre
YL = income of occupants of houses built on lot .3–.5 acre
YRT = average income of all rental occupants
YR3 = average income of rental occupants in structures built pre-1939
YR4 = average income of rental occupants in structures built 1940–49
YR5 = average income of rental occupants in structures built 1950–59
YR6 = average income of rental occupants in structures built 1960–65

P = matrix of housing prices

P3 = price of 5.5-room house built pre-1939, on .2–.3 acre lot, in sound condition
P4 = price of 5.5-room house built 1940–49 on .2–.3 acre lot, in sound condition
P5 = price of 5.5-room house built 1950–59 on .2–.3 acre lot, in sound condition
P6 = price of 5.5-room house built 1960–65 on .2–.3 acre lot, in sound condition
PS = price of 5.5-room house built 1950–59 on .2 acre lot, in sound condition
PL = price of 5.5-room house built 1950–59 on .3–.5 acre lot, in sound condition
R3 = monthly rent for 4-room apartment in structure built pre-1939
R4 = monthly rent for 4-room apartment in structure built 1940–49
R5 = monthly rent for 4-room apartment in structure built 1950–59
R6 = monthly rent for 4-room apartment in structure built 1960–65

S = matrix describing stock availability

OWN = percent of total stock which is single-family structures
S3 = owner-occupied units built pre-1939 as percent of all owner units
S4 = owner-occupied units built 1940–49 as percent of all owner units
S5 = owner-occupied units built 1950–59 as percent of all owner units
S6 = owner-occupied units built 1960–65 as percent of all owner units
SS = owner-occupied units built on .2 acre as percent of all owner units
SL = owner-occupied units built on .3–.5 acre as percent of all owner units
SR3 = rental units built pre-1939 as percent of all rental units
SR4 = rental units built 1940–49 as percent of all rental units
SR5 = rental units built 1950–59 as percent of all rental units
SR6 = rental units built 1960–65 as percent of all rental units

TABLE 6.3
STRUCTURAL EQUATIONS: RENTERS

Equation Number															
8	*YRT* =	2959	+51.26	*RT*	+17.83	*ED*	−3.885	*EA*	−149.1	*AGR*	+805.6	*R1*	+411.9	*R2*	R^2 = .6576
		(3.09)	(3.76)		(.76)		(.31)		(3.77)		(1.88)		(.76)		
9	*YR3* =	1233	+29.41	*R3*	+67.50	*ED*	−27.72	*EA*			+840.5	*R1*	+490.3	*R2*	R^2 = .3990
		(.69)	(1.79)		(2.65)		(1.36)				(1.08)		(.50)		
10	*YR4* =	263.9	+77.79	*R4*	+4.74	*ED*	+2.94	*EA*			+173.7	*R1*	+169.1	*R2*	R^2 = .3100
		(.11)	(3.91)		(.15)		(.11)				(.18)		(.14)		
11	*YR5* =	+1308	+50.85	*R5*	+35.21	*ED*	−13.35	*EA*			+1917	*R1*	+708.5	*R2*	R^2 = .3328
		(.65)	(3.04)		(1.18)		(.57)				(2.23)		(.66)		
12	*YR6* =	+2391	+41.17	*R6*	+41.41	*ED*	−33.91	*EA*			+1087	*R1*	+25.7	*R2*	R^2 = .4658
		(1.21)	(2.91)		(1.35)		(1.41)				(1.24)		(.02)		
20	*RT* =	−7.66	.010	*YRT*	−2002	*OWN*	+.2025	*EA*			−10.15	*R1*	−6.82	*R2*	R^2 = .3885
		(.48)	(3.45)		(1.45)		(.75)				(1.65)		(1.02)		
21	*R3* =	6.73	+.0069	*YR3*	−.2121	*OWN*	+.2366	*EA*	+.3164	*SR3*	−9.32	*R1*	−8.16	*R2*	R^2 = .3063
		(1.50)	(1.38)		(1.01)		(70)		(1.65)		(.95)		(.69)		
22	*R4* =	−1.73	+.0103	*YR4*	−.1893	*OWN*	+.0251	*EA*	+.0590	*SR4*	−2.50	*R1*	−3.42	*R2*	R^2 = .2553
		(.49)	(3.95)		(1.17)		(.88)		(.10)		(.22)		(.26)		
23	*R5* =	13.01	+.0098	*YR5*	−.2621	*OWN*	+.0828	*EA*	−.2050	*SR5*	−19.02	*R1*	−16.75	*R2*	R^2 = .2056
		(1.67)	(3.51)		(1.72)		(.32)		(1.64)		(1.75)		(1.44)		
24	*R6* =	15.88	+.0130	*YR6*	−.0700	*OWN*	+.3099	*EA*	−.5990	*SR6*	−16.15	*R1*	−5.78	*R2*	R^2 = .2297
		(.25)	(3.22)		(.42)		(.86)		(2.36)		(1.37)		(.42)		

means of the independent variables for average owner and renter incomes and prices. As expected, of the several determinants of the level of average neighborhood income, parents' education exhibits the highest elasticity: .882 for owners and .339 for renters. Elasticities of owner income with respect to lot size and structure age are .133 and .291, respectively. The availability of housing as reflected by the coefficient for the percentage of single-family structures has an elasticity of .408. Most striking is the low elasticity of owner income with respect to employment access, −.071. The coefficient of variation for structure age and tenure mix in this sample is two and one-half to three, and about one and one-half for lot size and employment access. Thus, next to parents' education, the most important determinants of variation in incomes across neighborhoods are the age distribution and tenure mix of the available housing.

Reduced-form coefficients for rent levels and incomes of rental occupants are different from those for owners in two respects, exhibiting higher elasticities with respect to structure age and lower elasticities with respect to parents' education.

In short, neighborhood income elasticities with respect to supply availability are less than price elasticities, while neighborhood income elasticities with respect to the characteristics of the available stock are higher than price elasticities; thus, variations in the available stock alter neighborhood incomes more than housing prices, whereas variations in supply affect prices more than incomes. This reflects the tendency of households to choose substitute housing types and locations as a function of prevailing prices.

Core Versus Suburban Income Differentials. The reduced-form equations reveal how income (and price) differences across geographic submarkets are related to differences in parents' education, racial composition, employment access, and housing-structure characteristics. The structure of the model implied by the reduced-form equations can be illustrated by relating the levels of neighborhood income and prices in selected submarkets predicted by the model to the actual levels of the independent variables in those submarkets. For purposes of illustration, the selected submarkets in which housing prices were described in some detail in Chapter 3 were chosen (see Table 3.4). (It is to be expected that reasonably good estimates will result, since these submarkets are from the sample used to estimate the model for prices and incomes.) The resultant estimates are shown in Table 6.5. The fact that incomes in the "nearby suburb" are almost as high as those in the core is attributable to a somewhat newer stock and better education. This shift in the demand price nearly offsets the negative effect of a greater supply

TABLE 6.4
REDUCED-FORM EQUATIONS

A. Coefficients for Owner Equations
Exogenous Variables

Endogenous	*EA*	*EDP*	*LT*	*AG*	*OWN*	*S3*	*S4*	*S5*	*S6*	*SS*	*SL*	*R1*	*R2*
YT	−59.4	959	9818	−201.8	−85.2							−2416	−2810
Y3	−61.2	983	10644		−46.5	−32.2						−5128	−6414
Y4	−54.0	544	5713		−27.3		−63.5					−1774	−1804
Y5	−150.4	1417	10792		−28.6			−136.9				−4014	−8926
Y6	−29.6	2094	72		−110.9				−190.2			−4654	−9242
YS	−39.6	986		−248.0	−95.8					−31.6		−1666	−3497
YL	−21.7	990		155.6	−207.6						−66.9	−613	−7298
PT	−110.3	1060	14797	−421.9	−489.9							−4868	−7965
P3	−91.9	912	9844		142.8	−99.7						−6814	−12734
P4	−224.1	1848	5122		100.6		−119.2					−6537	−9065
P5	−250.2	1591	12085		−69.2			−333.0				−6516	−13600
P6	−43.2	2253	80		−332.9				−570.6			−5073	−10495
PS	−50.6	2986		−719.2	−557.9					−183.8		−1188	−7906
PL	−51.6	1287		−201.6	−639.0						−206.4	5084	10392
ED		6.66										−10.8	−14.5

B. Coefficients for Renter Equations
Exogenous Variables

Endogenous	*EA*	*EDP*	*AGR*	*OWN*	*SR3*	*SR4*	*SR5*	*SR6*	*R1*	*R2*
YRT	−29.31	242	−304.2	−20.9					188	−400
YR3	−25.96	561		−7.8	11.6				−206	−1035
YR4	24.40	155		735		23.0			−360	−830
YR5	−34.70	459		−26.1			−20.6		1132	−1291
YR6	−100.27	587		−6.2				−52.9	−101	−1745
RT	373	2.41	−3.04	−.408					−8.14	−10.79
R3	−.051	3.74		−.265	.395				−7.22	−12.25
R4	−.020	1.30		−945		.295			−6.00	−12.15
R5	−.099	4.61		−.519			−.406		−7.15	−29.18
R6	−.064	5.87		−.150				−1.287	−20.85	−24.66

C. Elasticities at Mean for Selected Variables

Owners	*EA*	*EDP*	*LT*	*AG*	*OWN*
YT	−.071	.882	.133	.291	−.408
PT	−.062	.405	.091	.276	−1.190
ED	-	1.271	-	-	-

Renters	*EA*	*EDP*	*AGR*	*OWN*
YRT	−.058	.339	.729	.161
RT	.048	.211	.476	.212

TABLE 6.5
SOURCE OF VARIATION IN INCOMES OF OWNER-OCCUPANTS BETWEEN SELECTED CORE AND SUBURBAN SUBMARKETS

			New Suburbs				
Submarket	Central-City, Old Housing, High Density[a]	Established Nearby Suburbs, Newer Housing, High Density[b]	Medium Density[c]	High Density[d]	Very Low Density, 85% Single-Family Structures[e]	Suburban Industrial Job Center, Older Stock, 40% Black[f]	Suburban Job Center, Newer Stock, Medium Density[g]
Submarket characteristics (defined by exogenous variables):							
EA	6.5	11.5	21.5	16.5	24.5	14.0	12.5
EDP	11.4	11.8	12.1	12.1	12.5	9.5	11.8
LT	.09	.09	.20	.12	.60	.10	.19
AG	27.0	24.2	15.0	12.0	12.0	23.3	9.4
OWN	15.0	60.0	65.0	75.0	85.0	44.0	76.0
Mean owner income:	$13,300	$12,031	$12,227	$10,275	$14,715	$9,606	$10,498
Income differential, core versus selected suburbs:		−$1,269	−$1,073	−$2,725	+$1,415	−$3,694	−$2,802

Predicted effects of submarket characteristics on income differentials relative to core, based on reduced-form coefficients:							
Difference in:	*EA*	−$297	−$650	−$594	−$1,069	−$445	−$356
	EDP	383	671	671	986	−1,822	383
	LT	0	1,079	294	5,007	98	982
	AG	565	2,421	3,027	3,027	746	3,551
	OWN	−4,334	−4,260	−5,112	−5,963	−2,469	−5,702
Race effect:		0	0	0	0	−2,416	0
Predicted income differential between core and suburbs:		−$3,138	−$739	−$1,714	+$1,988	−$4,662	−$738

[a] San Francisco, Tracts A-1 to C-1, J-2 to J-20.
[b] San Francisco, Tracts 0-1 to 0-9.
[c] San Mateo County, Tracts 60 to 68.
[d] San Mateo County, Tracts 16 to 21.
[e] Contra Costa County, Tracts 38 to 55.
[f] Contra Costa County, Tracts 76 to 82 (in city of Richmond).
[g] Contra Costa County, Tracts 27 to 37 (in city of Concord).

of owner units. The negative effect on incomes of worse job access vis-à-vis the core is empirically quite small. In comparisons among suburbs outside the central city, the principal explanations for income variation lie in the characteristics of the stock. New units and low density tend to raise incomes. Table 6.5 includes three suburban areas largely distinguished by lot size. Lot-size effects on income are substantial. The model predicts an average of $11,186 in the newest suburb (average age of housing twelve years) with .12 acre lots versus $14,788 for a suburb slightly farther away but with an average lot size of .60 acre. (As would be expected, these estimates closely approximate actual income levels in suburbs with these characteristics in the Bay Area.)

The Effects of Race. Several competing hypotheses have been offered for the low incomes and low housing prices prevailing in central-city ghettos. As has been noted previously, ghetto areas tend to be nearer employment but have an older housing stock, higher densities, and a higher proportion of rental units. Assuming a perfectly competitive housing market, an old, primarily rental, and low-quality housing stock in the ghetto should exhibit lower incomes and prices, since black households are substantially poorer than whites. However, if discrimination is important, race may exert an independent effect on incomes and prices, independent of the simple demand-supply model of price and income determination described above. The captive demand by blacks may raise prices for certain types of ghetto housing.

Disentangling the separate effects of race has proven elusive. While many comparisons of incomes and prices inside and outside ghettos based on Census tract or other aggregate data have been made, these comparisons do not provide a satisfactory explanation of the cause of the differences, since they do not take into account differences in the housing stock, access, and other neighborhood characteristics. By including race as an additional variable in the price and income equations, the model can test whether race effects incomes and prices *independent* of such variables as differences in the quality of the housing stock, the supply available, and employment access. The race dummies are the net effect of all adjustments made by black and white households in the context of a discriminatory market.

The estimated reduced-form coefficients reveal that both prices and incomes are lower in black neighborhoods. Race reduces average owner incomes in the ghetto more than rental incomes. Incomes in owner-occupied units in ghetto submarkets are lowered by $2,810, while renters' incomes in ghetto submarkets are lower by $400 (reductions of 29.6 percent and 5.5 percent, respectively) from expected levels based on existing housing stock and neighborhood characteristics and assuming white occupancy. That incomes are significantly lower in ghetto submarkets, independent of housing stock and accessibility variables, reflects the pattern of out-migration of blacks (and whites) in the face of housing-market discrimination; for both white and black households, the likelihood of a household choosing to reside in a black submarket decreases as income rises.

Race also exerts an effect on prices in the black submarket. The reduced-form coefficients for race reveal that prices for owner-occupied units built prior to 1939 (over three-fourths of the owner stock in the core) are predicted to be lower by $6,414 and $5,128, respectively, in all black and mixed-racial areas, due solely to race. These are reductions of 21.9 percent and 14.4 percent from levels which would prevail *given* the type of housing and education and assuming white occupancy. Thus, the separate effect of racial composition is to reduce incomes in the black market *more* than it reduces prices for owner-occupied units, supporting the hypothesis of a shortage of owner-occupied units for black vis-à-vis white submarkets. In contrast, race reduces rents in the ghetto by $10.79, or 10.9 percent, and in mixed-racial areas by $8.14, or 7.7 percent, slightly more than the percentage reductions in renters' incomes.

The reduced-form coefficients for race also reveal that a neighborhood's being black implies a different pattern of incomes and prices *across* housing-quality classes (see Table 6.4). Among owner-occupied units, the race dummies imply that race reduces the variation in income across housing-age types and lot sizes (in comparison to comparable white submarkets). The explanation probably lies in the discrimination barriers which limit the entry of middle-class blacks to white submarkets. Though some middle-class blacks obtain entry to white submarkets, many others remain in the black submarket. As a result of the shortage of higher-quality units available in the black submarket, many of these middle- and upper-income blacks are forced to live in older, high-density units in the ghetto. In contrast, the race effect *increases* the price premiums for quality (both with respect to lot size and age of structure) across owner-occupied structures, presumably reflecting the substantial competition for the scarce newer units and those with larger lots available. There is no evidence of unusual price competition among rental units in the ghetto; income and price spreads of different quality levels are not significantly different in the black than in the white submarket.

In explaining actual observed differences in incomes and prices between ghetto and nonghetto submarkets, the quality and availability of the housing stock, parental education, and employment accessibility must also be considered, along with race. In San Francisco, ghetto areas tend to be nearer employment but have an older stock, higher densities,

TABLE 6.6
SOURCES OF VARIATION IN INCOMES OF OWNER-OCCUPANTS BETWEEN BLACK SUBMARKETS AND SELECTED WHITE SUBMARKETS

	Oakland Ghetto[a]	Central White Areas, Contiguous to Ghetto[b]	Established Nearby Suburbs, High Density[c]	New Suburb, Very Low Density, 85% Owner Occupancy[d]
Independent variable describing submarket:				
EA	6.0	6.0	12.0	24.5
EDP	9.5	12.1	10.7	13.3
LT	.09	.11	.13	.60
AG	28.0	25.2	14.6	12.0
OWN	30.0	35.0	68.4	85.0
Mean income:	$7,405	$11,950	$11,494	$14,715
Income differential between black and white submarkets:		$4,545	$4,089	$7,310
Predicted effects of submarket characteristics on income differentials between ghetto and white submarkets:				
EA		$0	$356	$1,099
EDP		−2,493	−1,151	−3,644
LT		−196	−393	−5,072
AG		−565	−1,502	−3,228
OWN		426	3,272	4,686
Total income effect		−$2,828	−$418	−$6,159
Race coefficient in reduced form		−$2,810	−$2,810	−$2,810
Total predicted income differential		−$5,638	−$3,228	−$8,969
Percentage breakdown of estimated income differential:				
Effect of neighborhood characteristics		50.2	12.9	68.7
Effect of race		49.8	87.1	31.3
Total		100	100	100

Note: Income and stock characteristics approximate prevailing levels in the various submarkets.

[a] Berkeley, Census Tracts BE-1A to BE-2C, Oakland, Census Tracts OK5A to OK24.

[b] Oakland, Census Tracts OK25-OK45B and Alameda, Tracts A69 to AL-16B

[c] Oakland, Census Tracts HA46 to HA55.

[d] Contra Costa County, Tracts 38 to 55.

and a higher proportion of rental units. Table 6.6 shows the effects of race versus these other characteristics of the housing stock and the neighborhood in explaining income differentials between selected black and white submarkets. The low quality of the housing stock much reduces ghetto incomes; the model predicts that incomes in ghetto areas would be below contiguous central areas by $2,828 due to these housing market differences *if* they were inhabited by whites. Significant differences would also exist between black submarkets and *suburban* submarkets if whites inhabited each. Actual observed income differentials between black and white submarkets are much larger, of course, the added differential essentially accounted for by the large race effect. Some of this could be accounted for by neighborhood-quality differences.

4. SUMMARY

This model explicitly recognizes the inherent compartmentalization in the housing market arising from the differences in the housing stock by location—differences which evolve very slowly. Models relying on one-way causation between income and prices either in their theoretical specification, or as a basis for econometric estimation, are directly at odds with this approach. While the specification described here does not provide a closed-form solution indicating how all markets are cleared, it does provide a consistent means for estimating how small changes in employment access or other housing submarket characteristics alter incomes and prices by housing type. A significant percentage of the spatial variation in incomes and prices is accounted for. The results reveal that housing-stock availability and housing-stock characteristics have a significant effect on incomes and prices. New, large-lot neighborhoods with limited entry by rental occupants are the areas which can be expected to exhibit the highest incomes and prices.