Housing Demand in the Developing-Country Metropolis

Gregory K. Ingram

A policymaker who realistically hopes to cope with urban housing problems should understand housing demand and supply conditions in the city. Effective policy requires knowledge of how markets allocate resources to housing, how homeowners and renters bid for dwellings, how developers and contractors respond to housing demand, and how government regulations and actions stimulate or constrain housing market activity. Our knowledge about many of these phenomena is modest, but our current understanding of housing markets in developing-country cities supports several policy generalizations.

The urban poor in developing and developed countries alike are the most likely to suffer from misallocated housing resources. Slum removal worsens housing conditions for the poor unless affordable low-cost dwellings are provided as substitutes for destroyed units. Often, new units must satisfy quality standards for space, structure, and services, which substantially add to construction costs. Unless cities underwrite large housing subsidies—which are beyond the fiscal means of most city budgets in developing countries—most new construction that meets high quality standards will be affordable only to middle- and high-income families. City managers must recognize that housing goals that are set too high translate into no housing for those who have the lowest incomes.

Note: The author has benefited from comments and discussions with Michael Hartley, Stephen Mayo, Janet Pack, Joseph de Salvo, Peter Schmidt, and participants in seminars at Princeton University, Michigan State, MIT, the World Bank, and the Corporación Centro Regional de Población. He thanks Sungyong Kang for research assistance, Maria Elena Edwards for manuscript preparation, and the staff of the Departamento Administrativo Nacional Estadística, Colombia, for aid with the data.

Developing countries can take several steps to more effectively meet the housing needs of their cities. Most important is the identification of those tasks best accomplished by municipal governments and those best left to private developers and service suppliers in the housing market.

Some services are most efficiently provided by the municipalities. Electricity, water, and sewerage utilities have substantial scale economies that public corporations may be best suited to provide. In addition, public health can be effectively protected only if everyone participates in proper waste disposal and disease prevention programs. Government services provided to all in such cases promote the public good. Local authorities, by establishing and enforcing laws, can also define property rights and facilitate the transfer of property. In developing countries customs regarding definition of property rights may hamper commercial and residential land use. In some developing countries unequal treatment under the law may bar groups from participating in urban development.

Many well-intentioned municipal policies should be avoided because of their perverse effects. Building codes inform tenants of housing quality, but they can be harmful if requirements are set too high. Zoning effectively bars developers from providing low-income housing in many sections of cities. Rent controls discourage prospective developers by reducing future return on housing investment, and existing structures fall into disrepair if the returns from controlled rents do not allow for maintenance expenditures. Similarly, usury laws that prevent banks from charging rates that cover their costs for small loans can effectively bar poor people from formal housing finance and leave them with the options of progressive construction or informal finance for housing. Municipalities can, however, sponsor credit programs when scale economies are involved and when private financial intermediation is primitive. Risk sharing can extend credit opportunities to those for whom such options previously did not exist.

Private developers should be encouraged to provide for the housing needs of all urban dwellers. With higher incomes from urban employment and with expanded credit opportunities, the urban poor can finance their own housing when standards are set appropriately. Some cities may decide to subsidize urban housing, which would enhance the resources at the command of the poor. Fundamentally, though, sound urban housing policy should be rooted in practices that are self-supporting if housing is not to be a drag on urban development.

The scope for analysis of housing issues is obviously broad, and this chapter addresses only one of many issues. It looks exclusively at urban housing demand and provides quantitative estimates of the factors that appear to determine housing demand. Such quantitative estimates are required to design policies that permit the supply of housing which is affordable for the urban poor.

Three sets of results related to the estimation of housing demand equations are reported. First, estimates of housing demand parameters based on household interview data from Bogotá and Cali, Colombia, are given. A comparison of parameter values with those obtained from North American data sets shows that the Colombian demand elasticities are generally comparable in magnitude to those from the United States. Second, an approach, derived from residential location theory, that stratifies people by place of work is employed to represent housing price variation in the demand equations. Finally, a simple exercise illustrates the magnitude of bias of the income elasticity of demand that can result from incorrect data aggregation techniques. Correctly aggregated data produce income elasticity estimates that are similar to those obtained from disaggregate or micro data.

The Price Term in Housing Demand Equations

Estimating demand equations for housing from cross-sectional data presents many challenges, but measuring the variation in the unit price of housing is probably one of the greatest difficulties. Data sets typically report the total expenditure on housing rather than a unit price and the quantity of housing. Hence the unit price must be inferred by relating variations in expenditure to variations in quantity. Moreover, housing is inherently multidimensional; it includes attributes of size, dwelling quality, location, public services, and neighborhood amenities that are obtained in a single, tied purchase. Since there is no widespread agreement as to how to measure the quantity of housing, one analyst's price variation may be another analyst's quantity variation. Finally, even if we can agree that housing prices may vary, it is not obvious that all price variation should be included in a housing demand equation. For example, if a metropolitan area's housing prices vary with the location of housing but households can locate anywhere, we cannot simply put the price actually paid by the household into the demand equation, because the household faces the whole schedule of prices. Simple inclusion of price indexes in a demand equation requires that households be in different market segments.

Numerous approaches have been employed to deal with one or more of these difficulties. Some examples follow.

1. Assume that intrametropolitan price variation does not exist, so that all variations in expenditures reflect variations in quantities; use expenditures in demand analysis as an index number to measure quantities (Muth 1969).
2. Allow intrametropolitan prices to vary across neighborhoods; estimate neighborhood-based price indices; then estimate demand equations on the assumption that residents of each neighborhood face only the prices in their own neighborhood (King 1975).
3. Allow intrametropolitan prices to vary by individual dwelling unit; estimate a dwelling unit price index with the use of a production function for housing and varying input prices; estimate demand equations on the assumption that occupants of each dwelling unit face only the price of their own dwelling unit (Polinsky and Elwood 1979).
4. Allow the marginal cost of attributes to differ within a metropolitan area; estimate a nonlinear hedonic price index and use the first derivative of the index with respect to specific attributes as the price term in a demand equation for the attribute (Witte, Sumka, and Erekson 1979).

Each approach has potential shortcomings. Omission of price variation, as in (1), can bias other demand equation parameters if the omitted price term is correlated with included variables. The assumption that households face only their neighborhood or dwelling unit prices, as in (2) and (3), may lead to a fundamental misstatement of the price variation in the sample if households are not limited in their choices to specific neighborhoods or dwelling units. If all purchasers face all prices, the price "chosen" may reflect the impact of other household characteristics. Price variations based on neighborhood or dwelling unit require that market segmentation based on those dimensions be justified. Estimation of demand equations for specific attributes of housing, as in (4), may not be relevant if we are really interested in the demand for housing as a composite good.

A relatively simple application of residential location theory suggests an alternative way of incorporating price variation into a demand equation for housing as a composite good. Simple models of residential location theory are essentially based on the precepts of cost minimization. A worker surveys the housing market from his workplace, j. He typically observes that housing prices, R, decline with distance, d, from his workplace in at least one direction, but that travel costs, t, increase with distance from his workplace. For any given amount of housing, H, he faces a total expenditure on housing, Z, composed of a housing expenditure plus a transport expenditure,

$$Z_j = R_j(d)H + t_j(d). \tag{11-1}$$

For quantity H_0 the worker can solve for the least-cost distance by taking derivatives

$$Z_j' = R_j'(d)H_0 + t_j'(d) = 0 \tag{11-2}$$

and solving the expression for d_j^*, the optimal distance or location for quantity H_0 and workplace j. This least-cost distance can be substituted back into equation 11-1 to calculate the minimum total expenditure for quantity H_0,

$$Z_j^* = R_j(d_j^*)H_0 + t_j(d_j^*). \tag{11-3}$$

Consider carrying out this exercise for different workplaces in a metropolitan area. The decline of housing prices with distance, $R_j(d)$, differs systematically across workplaces and probably shows steep rates of decline with distance for centrally located workplaces and gradual rates of decline for peripheral workplaces. Travel costs per unit distance may also differ by workplace, but in ways that may be difficult to generalize. For example, transit speeds may be higher but transit headways longer at peripheral locations as compared with a central location. As the workplace varies, however, so does the optimal housing and travel expenditure required for housing quantity H_0. This variation in expense by workplace for a given quantity of housing is used as a measure of price variation in the housing demand equations estimated here. A price index is estimated for each workplace zone. Households whose heads work at a particular workplace zone face the same housing price index, households with heads at another workplace face the price index at their workplace, and so forth. Prices will vary by workplaces.

If housing prices vary by workplace, it is worth asking why all workers do not try to obtain jobs at the workplace that has the lowest housing price index. Urban economists have long argued that a metropolitan area with multiple workplaces and a price gradient for housing must have differential wage levels across workplaces for households to be in equilibrium (Moses 1972). Accordingly, workplaces can have different housing prices, but they then must also have compensating differentials in wages to keep households in equilibrium. The existence of wage gradients across workplaces thus becomes a necessary condition for the workplace-based housing price variation approach taken here.[1]

Housing Demand and Workplace-Based Price Variation

In developing a workplace-based price index for housing, there are two possible formulations for the demand system, each of which uses a different definition of the price of housing. Differing definitions alter the specification of the demand equations. In one formulation the price of housing is based only on the housing expenditure and does not include the travel expenditure. In this case the budget constraint is written

$$Y = P_H H + P_v V + t(d) \tag{11-4}$$

where Y is income, P_H the price of housing, and P_v the price of composite commodity V. In this formulation the travel expenditure, $t(d)$, is included in the income constraint, and the derived demand equation is

$$H = f\{P_H, [Y - t(d)]\}. \tag{11-5}$$

That is, travel costs have to be subtracted from income in the demand equation. If travel costs are an unknown function of distance, d, then d is included in the demand equation as a separate variable.

In the second possible formulation the price of housing is the so-called gross price and is based on the housing expenditure plus the travel expenditure. In this case the budget constraint is

$$Y = Z_H H + P_v V \tag{11-6}$$

where Z_H is the gross price term. The travel cost does not enter separately into the budget constraint, and the distance term will not appear in the demand equations. To implement this second approach, however, one must be able to specify a priori the travel cost function, which is a combination of out-of-pocket cost and the opportunity cost of travel time. Since insufficient information is available for Bogotá and Cali to permit specification of the travel cost function with confidence, the first approach has been implemented here. The estimated demand equations therefore include distance to the workplace, as in equation 11-5, and the workplace-based price term is based on housing expenditure only.

The relevant housing expenditure that is used to define a price index for a given workplace is the efficient, or optimal, expenditure implicit in the solution of equations 11-2 and 11-3. Corresponding to each quantity of housing, H, is an optimal location or optimal distance, d^*, and an optimal expenditure, $R(d^*)H$. If households employ the kind of locational calculus embodied in residential location theory, the choice made by households with a head employed at a particular workplace is at or near the optimal location for that workplace, and their housing expenditure approximates the optimal expenditure for their workplace and housing quantity.[2] The relation between housing expenditures and housing quantity for a given workplace can be captured by regressing the observed housing expenditure on measures of housing quantity for households whose heads work in the same work zone. The relation between housing expenditure and housing quantity can then be used to formulate a price index for the given workplace. This procedure can be repeated to yield a price index for each workplace, and these workplace-specific price indexes can then be used as a price term in demand equations for housing as a composite good.

The specific procedure that has been implemented in this chapter can be summarized as follows. The sample contains M households whose household heads have jobs located at one of J work zones, and there are N_j households associated with workplace j. For household i ($i = 1$ to N_j) at workplace j, the monthly expenditure on housing (or the dwelling unit value), R_{ij}, and a set of K dwelling unit characteristics, X_{ijk}, are known. For each of the J work zones an estimation is made of the equation

$$R_{ij} = \sum_{k=1}^{K} B_{jk} X_{ijk} \tag{11-7}$$

by regressing housing expenditure on the measure of dwelling characteristics, and J sets of parameters which indicate how the cost of housing attributes varies by workplace are obtained. A representative dwelling unit in the housing market is then defined as the unit that has the samplewide average amount of each dwelling unit characteristic, where the average quantity is

$$\bar{X}_k = \frac{1}{M} \sum_{j=1}^{J} \sum_{i=1}^{N_j} X_{ijk}. \tag{11-8}$$

The dwelling with attributes $\bar{X}_k$ then becomes the standard unit—the equivalent for housing of the standardized market basket. For each workplace the estimated parameters in equation 11-7 are used to calculate the cost of the standard unit,

$$\bar{R}_j = \sum_k B_{jk} \bar{X}_k. \tag{11-9}$$

The cost of a standardized unit is used to formulate a workplace price index by choosing workplace 1 as a numeraire and calculating a price index

$$\Pi_j = \frac{\bar{R}_j}{\bar{R}_1}. \tag{11-10}$$

The households in the sample also have associated with them C household characteristics, HC_c, that affect household demand for housing. These characteristics of the households and the distance from home to work, d_{ij}, are used in a demand equation whose dependent variable is housing expenditure divided by the price index in equation 11-10, or a quantity index of housing. The demand equation is of the form

$$\frac{R_{ij}}{\Pi_j} = f(\Pi_j, HC_{ijc}\, d_{ij}) \tag{11-11}$$

and is estimated over the sample of all M households as a single pooled demand function. In equation 11-11 both linear and double log specifications are used.

The Setting and the Data

The household interview data used to implement the housing demand procedure just outlined are from Bogotá and Cali, Colombia. The principal data set used was collected in 1978 and covers both owners and renters, for whom equations are estimated separately, in Bogotá and in Cali. A second data set is available for Bogotá in 1972, but data for renters only can be used to estimate the demand for housing in 1972.

In 1978 Bogotá had a population of roughly 3.5 million and Cali a population of roughly 1.1 million. Both cities have experienced rapid rates of population growth in the past (Bogotá's population in 1972 was 2.8 million), but current population growth rates are moderating in both cities. Per capita income in 1978 was about $800 a year in both places. The cities differ significantly in climate because of their different altitudes: Bogotá is 8,000 feet above sea level and has temperate weather with cool nights; Cali, at 3,000 feet above sea level, is semitropical and warmer than Bogotá. Differences in size and climate may explain some of the differences in housing demand in the two cities.

To implement the workplace-derived price indexes it was necessary to divide the two cities into a number of work zones the boundaries of which are arbitrary but are based on considerations that include compactness, respect for significant internal divisions, and a requirement that there be an adequate number of observations in each work zone. The same work zone system was used in Bogotá in 1972 and 1978 and for renters and owners in each city. Tabulations of residence and workplace by annular ring and radial sector indicate a high degree of association between place of work of the household head and place of residence of the household. Empirical analyses indicate that although the workplace of secondary workers may have a slight influence on a household's residential location, the workplace of the household head is clearly a dominant determinant (Pineda 1981). Average commute lengths in kilometers for each work zone and tenure type in both cities differ by up to a factor of 3. In both cities commute lengths are long for centrally located workplaces and for workplaces located along the mountains.

The Hedonic Price Equations

Separate hedonic equations were estimated for each work zone and tenure type in Bogotá and Cali in 1978. For 1972 in Bogotá a hedonic equation was estimated only for renters because no data were available for the value of owner-occupied units in the 1972 sample. For renters the dependent variable is the monthly rent; for owners the dependent variable is the value of the dwelling unit in thousands of pesos.

Because the 1978 data were all collected in the same survey with the same questionnaire, it is possible to use the same specification for the four sets of 1978 equations. In the 1978 equations the independent variables included the dwelling unit area in square meters, *DUAREA*, the lot area in square meters, *LOTAREA*, the number of blocks to the nearest bus line, *BLKTOBUS*, a dummy variable equal to 1 if the residence had a private or public phone, *DPHONACSS*, a dummy variable equal to 1 if the dwelling unit had its own nonshared kitchen and bathroom facilities, *DEXCLUSE*, and a dummy variable equal to 1 if the dwelling unit had its garbage picked up by municipal authorities, *DGARBCOL*. It is interesting to note the similarities and differences between tenure classes and cities in the independent variables. Renters in Bogotá and Cali, for example, have similar-size units on similar-size lots, but Bogotá renters have more phones, while Cali renters have better garbage collection. Bogotá owners have larger, more expensive homes on larger lots than do Cali owners. The most striking differences between renters and owners are in the average area of the unit and the proportion of units having exclusive bath and kitchen facilities; owners are better housed than renters. Finally, there is more variability in the average dependent variable across work zones than there seems to be in the average independent variables, which indicates that prices do vary.

The independent variables used in the 1972 equations differ from those used in 1978 because the questionnaire was quite different. The variables are building age, number of rooms, presence of garbage collection, distance to nearest bus line, type of structure, condition of unit, presence of nonresidential use, and presence of a public or private phone. Current prices are used in both time periods. The consumer price index had approximately tripled, from 47 in 1972 to 150 in 1978.

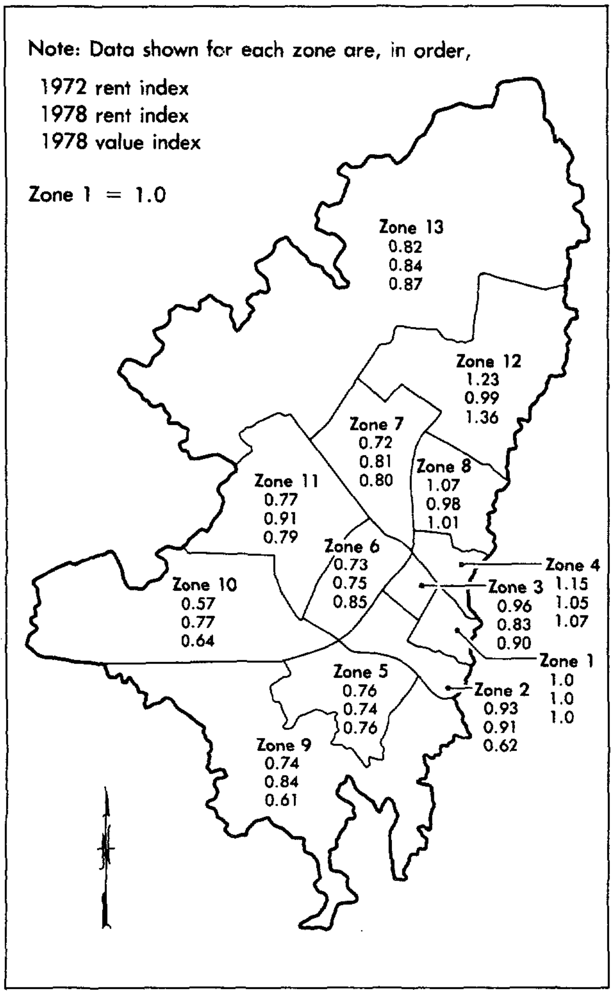
The coefficients from the hedonic price equations are most comparable for 1978. In 1978 there are equations for thirteen Bogotá and eight Cali work zones and for two tenure types, a total of forty-two equations. The only variable that always has the correct sign in all forty-two equations is dwelling unit area. Access to a phone, exclusive bath and kitchen facilities, and garbage collection also perform well; they have the expected sign thirty-six, thirty-seven, and thirty-two times, respectively. The number of blocks to a bus is positive only half of the time; it is possible that there is some disamenity associated with being too close to a bus route. Lot area does not perform well in the hedonic equations, and it does very poorly in Cali, where owners, in particular, do not seem to value larger lot size. The hedonic equations for the 1972 Bogotá renters are similar to those for 1978 in that the measure of interior space, the number of rooms, has a positive effect on rent.[3]

A measure of the explanatory power of the hedonic price equations is shown in table 11-1, which summarizes the explanatory power of the regression equations and the workplace stratification in an analysis of variance framework. Overall, the analysis explains from 45 to 69 percent of the variation in housing prices; the equations have much more explanatory power than does the workplace stratification. Interestingly, the workplace stratification has more explanatory power for owner-occupied units than for renter-occupied units. This is consistent with the empirical regularity that owner-occupied units have steeper price gradients in urban areas than do renter-occupied units. Hence, workplace location matters more in the owner market than in the renter market.

Table 11-1. ***Analysis of Variance: Hedonic Price Equations***

	Percent of variation explained by		
Data	*Work zone stratification*	*Equations*	*Total*
Bogotá			
1972 renters	4.7	49.3	54.0
1978 renters	2.5	47.6	50.1
1978 owners	8.7	36.4	45.1
Cali			
1978 renters	1.9	64.3	66.2
1978 owners	8.0	60.9	68.9

Map 11-1. *Bogota: Workplace Price Indexes, Thirteen Work Zones, 1972 and 1976*



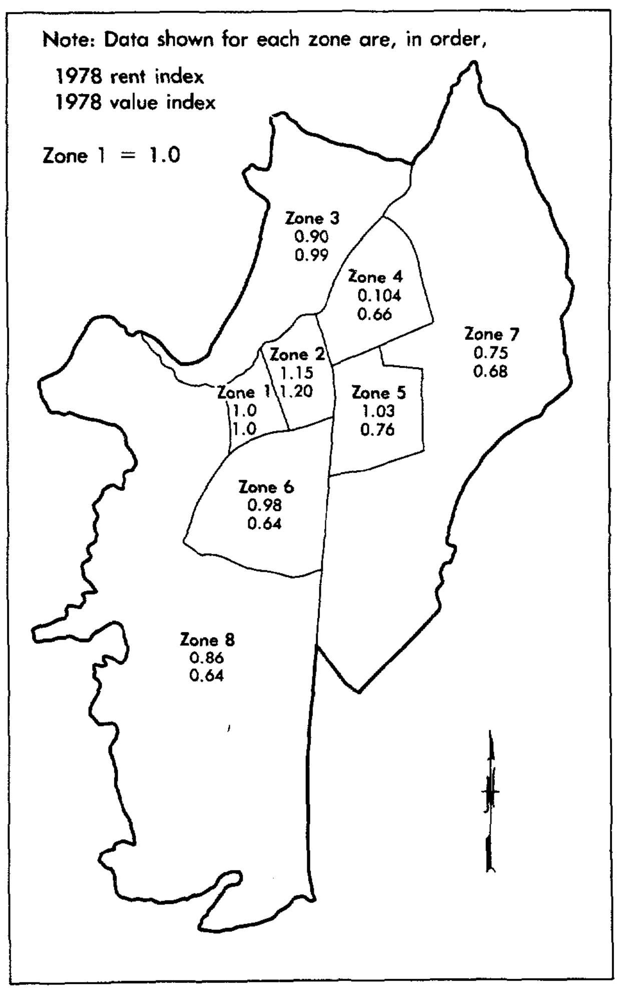
The standardized rents and values were obtained by plugging the average renter and owner unit characteristics for Bogotá and Cali into their respective workplace hedonic equations. For use in the demand equations these rents and values are transformed into spatial price indexes by dividing through by the relevant rent or value for work zone 1, the central business district. The resulting normalized price indexes are displayed for Bogotá in map 11-1 and for Cali in map 11-2. There obviously is variation across work zones in these price indexes. In both Bogotá and Cali there is more variation in the price index for owners (the range covers a factor of 2) than for renters.

The Housing Demand Equations

The dependent variable in the demand equations is the monthly rent or value divided by the workplace-specific price index, as shown in equation 11-11. The independent variables are monthly household income (a measure of current income) in pesos, the price index described above, and the air distance from home to work, in meters. Three additional household characteristics are included in the demand equations: a dummy variable for the sex of the household head (1 = male); family size (the number of persons in the household); and the age of the household head. The hypothesis is that these three characteristics capture differences in taste (sex of household head), in the need for housing (family size), and in assets or wealth (age of the household head). Two functional forms, double log and linear, are estimated. In the linear specifications squared terms for family size and the age of the head are entered to capture nonlinearity in the effects of those variables.

Five sets of equations are estimated; each set is a different combination of year, tenure choice, and city. A comparison of mean values across the five samples shows that renters have younger household heads, smaller families, and lower incomes than owners. Differ-

Map 11-2. *Cali: Workplace Price Indexes, Eight Work Zones, 1978*



ences between Bogotá and Cali are slight except for income; Bogotá owners have much higher average incomes than do Cali owners, and Bogotá renters have average incomes similar to those of Cali renters. 1978 Bogotá renters had smaller families and younger household heads than did 1972 Bogotá renters.

The demand equations perform well with R^2 statistics that range from 0.25 to 0.6. Income is by far the most important explanatory variable. Age of the head and family size are usually significant; sex of the head is usually not significant, but it always has a negative sign. The housing price index is significant in two of the five samples, and it always has the correct sign. Distance from home to work is significant in four of the five samples and has the correct sign in nine of the ten equations.

A summary of results from the fully specified demand equations, in the form of elasticities for each independent variable, is displayed in table 11-2 for renters and table 11-3 for owners. The elasticities are calculated in the linear equations with the use of the mean value of each independent variable except income. The linear elasticities are shown for the first, second, and third quartiles of each sample's income distribution. In each case the sample mean and the seventy-fifth percentile of the income distribution are essentially identical.

The magnitudes of the various elasticities obviously vary among the samples and specifications shown, but they display a consistent and stable pattern for most of the variables. All income elasticities are less than one, and at the sample mean they lie in a narrow range of 0.6 to 0.8 except for the equations for Cali renters. The elasticity of the sex of the head is always negative and small (in the interval 0 to −0.2). Family size elasticities show an interesting pattern; they are negative for owners and are usually positive for renters. Since renter-occupied units are usually smaller than owner-occupied units, it appears that space is a binding constraint for renters, and larger renter families obtain more housing. Owner-occupants seem to be able to reduce the quantity of housing demand as family size increases because they have larger units on the average and the quantity of housing is not constrained by family size. Age of the household head has a consistently positive demand elasticity when evaluated at the sample mean. By using the linear demand equation with the squared term for age of the head, it is possible to calculate the age at which housing demand is at a maximum. This is consistently ages 50–57 except for Bogotá owners, for whom it increases throughout the relevant range. The price elasticity of demand lies consistently in the interval 0 to −1.0 in absolute value and becomes absolutely quite small for some of the linear specifications. Finally, the distance elasticity is almost consistently negative and quite small.

Table 11-2. *Demand Elasticities at Various Income Levels, Renters*

Income percentile and level[a]	*Income*	*Sex of head*	*Family size*	*Age of head*	*Price*	*Home-to-work distance*
1972, linear, Bogotá						
25 (1,000)	0.32	−0.16	0.30	0.23	−0.91	−0.05
50 (1,700)	0.45	−0.13	0.25	0.19	−0.75	−0.04
75 (3,079)[b]	0.59	−0.09	0.18	0.14	−0.55	−0.04
1972, log-log, Bogotá						
All	0.77	−0.14	0.14	0.12	−0.70	−0.06
1978, linear, Bogotá						
25 (3,500)	0.55	−0.03	−0.24	0.95	−0.17	−0.23
50 (7,100)	0.71	−0.02	−0.16	0.61	−0.11	−0.15
75 (11,260)[b]	0.80	−0.003	−0.11	0.43	−0.08	−0.10
1978, log-log, Bogotá						
All	0.72	−0.07	0.10	0.07	−0.28	−0.06
1978, linear, Cali						
25 (3,500)	0.05	−0.01	0.48	0.62	−0.34	−0.16
50 (7,300)	0.10	−0.01	0.46	0.59	−0.32	−0.15
75 (12,829)[b]	0.16	−0.01	0.42	0.55	−0.30	−0.14
1978, log-log, Cali						
All	0.47	−0.20	0.36	0.43	−0.48	−0.03

a. Income levels are in parentheses.
b. Sample mean.

Table 11-3. *Demand Elasticities at Various Income Levels, Owners*

Income percentile and level[a]	*Income*	*Sex of head*	*Family size*	*Age of head*	*Price*	*Home-to-work distance*
1978, linear, Bogotá						
25 (6,000)	0.33	−0.03	−0.34	0.66	−0.31	0.02
50 (10,900)	0.47	−0.02	−0.27	0.52	−0.24	0.01
75 (17,942)[b]	0.60	−0.02	−0.21	0.40	−0.19	0.01
1978, log-log, Bogotá						
All	0.78	−0.09	−0.25	0.25	−0.44	−0.02
1978, linear, Cali						
25 (5,000)	0.39	−0.06	−0.57	0.53	−0.27	−0.06
50 (8,800)	0.53	−0.05	−0.44	0.18	−0.21	−0.05
75 (13,841)	0.64	−0.04	−0.34	0.13	−0.16	−0.04
1978, log-log, Cali						
All	0.76	−0.06	−0.30	0.08	−0.33	−0.02

a. Income levels are in parentheses.
b. Sample mean.

Table 11-4. *Range of Housing Demand Elasticities in Various Countries (from Household Observations)*

Country	*Elasticity of housing demand with respect to*				
	Current income	*Price*	*Family size*	*Age of head*	*Sex of head (1 = male)*
Renters					
Colombia	0.2 to 0.8	−0.1 to −0.7	−0.1 to 0.4	0.1 to 0.6	−0.01 to −0.2
United States	0.1 to 0.4	−0.2 to −0.7	?	?	Consistently negative
Korea, Rep. of	0.12	−0.06 to 0.03	0.15 to 0.25	—	—
Owners					
Colombia	0.6 to 0.8	−0.15 to −0.40	−0.2 to −0.35	0.1 to 0.4	−0.02 to −0.1
United States	0.2 to 0.5	−0.5 to −0.6	?	?	Negative
Korea, Rep. of	0.21	−0.05 to 0.07	−0.02 to 0.15	—	—

? Not conclusive.
— Not reported.
Sources: United States: Mayo (1981); Korea: Follain, Lim, and Renaud (1980).

Table 11-4 summarizes the range of demand elasticities obtained in Cali and Bogotá and compares them with estimates obtained from household surveys in the United States and Korea. The general pattern of results is quite similar for the United States and Colombia; both countries differ somewhat from Korea. The Colombian income elasticities are somewhat higher than those obtained in the United States, whereas the Colombian price elasticities may be lower. The elasticities of housing demand with respect to family size and age of the head in Colombia cannot be compared with data from the United States but are somewhat similar to the Korean estimates. Finally, the effect of the sex of the household head, although usually statistically insignificant in Colombia, is also always negative, as in the United States. There are three possible explanations for this result. First, female-headed households may have stronger preferences for housing than male-headed households. Second, female-headed households may be discriminated against and face higher prices, which could produce larger expenditures on housing, and those larger expenditures could show up as a preference for larger quantities in the demand equations for renters. The discrimination hypothesis is, however, unconvincing for owner-occupants. Third, female household heads have shorter commuting distances than male household heads and may therefore systematically pay higher prices for housing, since rents decrease with distance from the center of the city. The demand equations used should account for this difference, however, because distance is included. Accordingly, the first explanation, based on preference differences, may be the most plausible.

To investigate the effect of distance on the coefficient for the sex of the head and to see how sensitive the other parameters are to both the price and distance terms, the housing demand equations were estimated without those terms. Omission of the price and distance terms tends to reduce the income coefficient very slightly, often only in the third significant digit. The family size effects are also affected only minimally by the omission of the two terms. The coefficients for sex and age of head do change considerably in percentage terms, however; this seems to be largely a consequence of the omission of the distance term. Female-headed households live closer to the head's workplace than do male-headed households, as do households with older heads compared with households with younger heads. In general, however, the parameter estimates for the included variables are stable with respect to the omission of the price and distance terms.

These exercises suggest that neither the housing prices as specified in these demand questions nor the distance from home to work are collinear with household income. Indeed, in Bogotá and Cali, as in many other cities, the use of micro data dramatically reduces problems of multicollinearity in the estimation of housing demand equations.

Aggregate Estimates of Income Elasticities

All of the parameter estimates that have been presented so far have been obtained from computer-based multivariate regressions that use individual households as observations. In many situations it may not be possible to gain access to individual household records because of confidentiality restrictions, while in other situations lack of time or of adequate computer facilities may make parameter estimation with micro data impossible. This section briefly investigates the adequacy of the parameter estimates that could be made from published aggregated data. The focus is on the estimation of the income elasticity of the demand for housing because that parameter is often of interest in both the design and the evaluation of housing programs, policies, and projects.

Each of the five samples analyzed above is now summarized in a matrix dimensioned by rent or value and by income. Eight income categories were defined for the 1978 data and nine for the 1972 data. The average rent or value was calculated for each income category; this average was then regressed on the midpoints of the income categories in a log-log specification using a hand-held calculator. The equations yielded by this exercise are shown in table 11-5, and the resulting income elasticities are compared with those from the disaggregated, fully specified equations in table 11-6. The aggregate estimates each differ by less than 20 percent from the disaggregated log-log estimates, and in four of five cases the aggregate log-log estimates lie between the linear and log-log disaggregated estimates. It is obvious that aggregate-based estimates of the income elasticities of the expenditure for housing could be a good approximation for the income elasticity of demand for housing in the samples used here.

It is important to note that the aggregate estimates obtained are sensitive to the way in which the underlying micro data are aggregated. An experiment that illustrates this was performed with the 1972 sample of renters. The sample was aggregated to the level of sixty-three zones for the city of Bogotá, and average rents and incomes were calculated for each zone. A hand-held calculator was then used to calculate a log-log regres-

Table 11-5. *Housing Demand Equations from Aggregate Data*

Sample	B_0	B_1	R^2
1972 phase II renter	2.92	0.71	0.99
1978 Bogotá renter	1.54	0.79	0.99
1978 Cali renter	12.38	0.55	0.97
1978 Bogotá owner	9.11	0.67	0.99
1978 Cali owner	7.81	0.66	0.97

Note: The equation is of the form $R = B_0 \cdot Y^{B_1}$, where R is rent and Y is income.

Table 11-6. *Aggregate and Disaggregated Income Elasticities of Housing Demand*

Sample and specification	*Aggregate*	*Disaggregated*
1972 Bogotá renter		
Log-log	0.71	0.77
Linear	—	0.59
1978 Bogotá renter		
Log-log	0.79	0.72
Linear	—	0.80
1978 Cali renter		
Log-log	0.55	0.47
Linear	—	0.16
1978 Bogotá owner		
Log-log	0.67	0.78
Linear	—	0.60
1978 Cali owner		
Log-log	0.66	0.76
Linear	—	0.64

— Not applicable.

sion of average zonal rent on average zonal income, using all sixty-three observations. The resulting income elasticity, 0.95, was substantially higher than the 0.71 estimate that was obtained with nine observations from the correctly aggregated sample. A second experiment was then run on the 1972 Bogotá data. For this experiment the data in the rent-income matrix were *incorrectly* aggregated by calculating the average income for each rent category and regressing the rent category midpoints on the mean incomes. This rent-stratified approach yielded an income elasticity estimate of 1.36, nearly twice the 0.71 obtained with the use of an income-stratified aggregation procedure. It is obvious that the aggregation bias in estimates of income elasticities can be very large, but that correctly aggregated data can give useful results.

Conclusion

This paper has described and implemented a two-step estimation procedure for incorporating price variation in the estimation of demand equations for housing, using household survey data from Bogotá and Cali, Colombia. The demand equations estimated with this procedure give significant results for the income elasticity of the demand for housing; estimates of the income elasticity generally lie in the upper end of the 0.2–0.8 range. Although the price term in the demand equations gave less significant results, the price elasticity of demand appears to be between 0 and −1. There is, however, greater uncertainty about the magnitude of the price elasticity than about the magnitude of the income elasticity. Other household characteristics involved in the demand equations have low demand elasticities, typically less than 0.5 in absolute magnitude. The age of the head has a positive elasticity over most of its range, while family size usually has a positive elasticity for renters and a negative elasticity for owners. The demand equations suggest that female-headed households consume more housing than male-headed households, but this result is rarely statistically significant. Distance from home to work is entered into the demand equations as an adjustment to income, but it undoubtedly also represents price variation within the workplace strata, which are used as the main representation of price variation. The distance elasticity is small—less than −0.2—and is almost always negative.

Comparisons of elasticity estimates with those obtained from U.S. data sets indicate that the range of the Colombian estimates generally overlaps the range of the U.S. estimates. This similarity of values may seem surprising at first but is much less so on reflection. Housing is a nontraded good, and its price is endogenous to the local economy and reflects, among other things, local income levels. Perhaps we should be more surprised at the similarities between Bogotá and Cali, which have markedly different climates.

Simple experiments involving the aggregation of the household survey data used to obtain micro data estimates suggest that income elasticity estimates on the basis of correctly aggregated data can be good proxies for estimates on the basis of fully specified models that use household observations. Estimates on the basis of incorrectly aggregated micro data, however, can produce estimates of the income elasticity of demand that are badly biased.

Notes

1. Preliminary empirical work indicates that a wage gradient with a peak in the central business district does exist in Bogotá.
2. Equation 11-3 can be solved for the expansion path of expenditures as the quantity of housing increases.
3. Parameter estimates are available from the author.

Bibliography

Follain, J., G. C. Lim, and B. Renaud. 1980. "The Demand for Housing in Developing Countries: The Case of Korea." *Journal of Urban Economics*, vol. 7, no. 3 (May), pp. 315–36.

King, Thomas. 1975. "The Demand for Housing: Integrating the Roles of Journey-to-Work, Neighborhood Quality, and Prices." In N. Terleckyj, ed., *Household Production and Consumption*, pp. 451–83. New York: Columbia University Press.

Mayo, Stephen K. 1981. "Theory and Estimation in the Economics of Housing Demand." *Journal of Urban Economics*, vol. 10, no. 1 (July), pp. 95–116.

Moses, Leon N. 1972. "Toward a Theory of Intra-Urban Wage Differentials and Their Influence on Travel Patterns." *Papers and Proceedings of the Regional Science Association*.

Muth, Richard F. 1969. *Cities and Housing*. Chicago: University of Chicago Press.

Pineda, José Fernando. 1981. "Residential Location Decisions of Multiple Worker Households in Bogotá, Colombia." Presented at annual meetings of the Eastern Economic Association, Philadelphia, Pa., April. Processed.

Polinsky, A. Mitchell, and David M. Elwood. 1979. "An Empirical Reconciliation of Micro and Grouped Estimates of the Demand for Housing." *Review of Economics and Statistics*, vol. 61, no. 2 (May), pp. 199–205.

Witte, Ann D., Howard J. Sumka, and O. Homer Erekson. 1979. "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets." *Econometrica*, vol. 47, no. 5 (September), pp. 1151–73.