Volume Title: Housing Markets and Racial Discrimination: A Microeconomic Analysis

Volume Author/Editor: John F. Kain and John M. Quigley

Volume Publisher: NBER

Volume ISBN: 0-870-14270-4

Volume URL: http://www.nber.org/books/kain75-1

Publication Date: 1975

Chapter Title: Appendix D: Weighted Expenditure Models

Chapter Author: John F. Kain, John M. Quigley

Chapter URL: http://www.nber.org/chapters/c3725

Chapter pages in book: (p. 324 - 335)

D

Weighted Expenditure Models

The analyses presented in this book are based primarily on a nonrandom sample of St. Louis city and St. Louis County households. The probability of a particular household's being included in these samples depends on residence location and, in the case of county households, on the type of structure occupied.

The largest difference in sampling rates is between St. Louis city and St. Louis County. On the average, a city resident was about eight times as likely to be sampled as a county resident. Approximately one out of every 150 dwelling units were sampled in the city, as contrasted with only one out of every 1,250 dwelling units in the county. The rates for county residents vary, depending on whether they resided in single-family or multifamily units, with the sampling rate for single-family units about two and one-half times as high as that of county multifamily units. In the city, dwelling units located in low-income neighborhoods were sampled at a higher rate than those located in high-income neighborhoods.[1]

Since black households are concentrated in the poorest neighborhoods, the oversampling of low-income neighborhoods provided much larger samples of black households than would have been obtained in a simple random sample. The extensive analysis of the effects of housing-market discrimination which is a central feature of this book would not have been possible if a random sample of the same size had been obtained.

In spite of the clear advantages provided by the stratified sample, there are disadvantages. The sample contains fewer suburban properties than would be desirable. In addition, the use of observations represent-

[1]There are actually four city sampling rates based on 1960 census-tract income levels. They vary from a 1-in-66 rate for dwelling units located in tracts with median family incomes below $3,000 to a 1-in-224 rate for units located in tracts with median family incomes greater than $6,000. A more detailed discussion of the sample is included in Chapter 4.

ing different numbers of households raises questions about how the individual observations should be weighted in estimating parameters. The most obvious alternatives are weighting each sample observation equally or weighting each observation by the inverse of the sample proportion. Nearly all of the analyses presented in this book, and all of those presented thus far, employ an equal weighting of the sample observations. If the behavioral relations are properly specified, the weighting should make little difference, even though weighting by the inverse of the sampling rates is more efficient from a statistical viewpoint. If, however, misspecification is present, the choice of weights may have a substantial effect on the parameter estimates.

We acknowledge the possibility that the parameter estimates may be affected by the weighting procedure in our discussion of the income elasticities obtained for St. Louis households. As has been noted, the income elasticities obtained from the simple models of housing expenditures are substantially lower than those obtained in most previous studies. We speculated that part of these differences might be due to the underrepresentation in the sample of households with high income elasticities of demand for housing services. The use of neighborhood income in defining sampling rates would presumably have produced this result even for households at the same income level. To test for this possibility, we estimated simple models of housing expenditure for owners and renters, weighting each observation by the inverse of the sample proportion. These estimates both enable us to evaluate the effects of weighting on the parameter estimates of the simple expenditure models and provide some more general indication of the effect of using weighted estimates.

Some insight into the effects of weighting is provided by the weighted and unweighted mean values of proportion black, income, rent, and value in Table D-1. The weighted rental sample has a much smaller mean proportion black, .24, than the unweighted rental sample, .45; a higher mean annual income, $6,753, than the unweighted rental sample, $5,395; and a higher mean rent, $80.75 per month, than the unweighted rental sample, $63.31 per month. The differences in mean income and house value between the weighted and unweighted owner samples are even larger.

Estimates of the simple expenditure models for owners of one-family units and for renters, using the inverse of the sampling proportion as weights, are shown in Tables D-2 and D-3. As an aid to the reader, the comparable unweighted estimates of each equation are reproduced as well. The most clear-cut result of this weighting is to provide even larger estimates of underconsumption of housing services by black households than are obtained from the unweighted regressions. The coefficient of

TABLE D-1
Mean Income, Rent, Value, and Proportion Black for Weighted and Unweighted Samples of St. Louis Households

	Renters		Owners	
	Unweighted	Weighted	Unweighted	Weighted
All				
Income	$5,395	$6,753	$8,618	$10,218
Proportion black	.45	.24	.18	.07
Rent (value)	$63.31	$80.75	$16,512	$18,973
Blacks				
Income	$4,262	$4,529	$6,420	$6,231
Rent (value)	$53.66	$57.62	$11,781	$11,519
Whites				
Income	$6,313	$7,446	$9,099	$10,522
Rent (value)	$71.14	$87.95	$17,547	$19,541

the race variable is larger in absolute value in the weighted than in the unweighted regression in all six equations (three specifications for renters and owners). For example, the unweighted additive rental equation indicates a difference in monthly housing expenditures of comparable black and white renters of $15.34, as compared to a difference of only $9.52 for the unweighted equation. Similarly, a coefficient of −.156 was obtained for the unweighted estimate of the semilog equation, and an estimate of −.216 for the same equation using sample weights. The estimated monthly expenditure of black owners based on the unweighted additive model is $37.64 per month ($3,764 in housing value) less than that of similar white owners; using weighted estimates of the same model, the difference is $48.07 per month ($4,807 in housing value). Comparable differences were obtained for the weighted and unweighted estimates of the semilog and log-log specifications of the owner model.

Weighting has a somewhat smaller effect on the income coefficients. In the additive renter model, the weighted estimates of the income coefficient are nearly a third larger than the nonweighted estimate, $2.85 versus $3.79 per thousand dollars of annual income. The elasticities, computed at sample means in the linear equation, increase from .24 for the unweighted model to .32 for the weighted model. Similarly, weighting for the log-log specification of the rental model increases the estimate of the income coefficient, the constant elasticity, from .14 to .18. For the semilog model, however, the coefficient of income is the same in both the weighted and unweighted equations.

TABLE D-2
Alternative Specifications of Simple Expenditure Models for Renters, Using Weighted and Unweighted Samples

	Linear		Semilog		Log-Log	
Variables	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
Race	−9.52[1]	−15.34[1]	−.156[1]	−.216[1]	−.187[1]	−.231[1]
Income	2.85[1]	3.79[1]	.045[1]	.045[1]	.137[1]	.182[1]
Corrections						
No furniture	8.90[2]	11.99[2]	.093[4]	.163[2]	.124	.241[1]
No heat	−12.63[1]	−6.22[2]	−.260[1]	−.163[1]	−.272[1]	−.126[1]
No appliances	−21.51[1]	−25.58[1]	−.196[1]	−.292[1]	−.216[1]	−.392[1]
No water	1.34	−8.41[1]	.004	.128[1]	.006	.112[1]
Constant	70.74[1]	69.44[1]	4.15[1]	4.19[1]	3.26[1]	2.93[1]
R^2	.38	.50	.33	.47	.28	.42
Income elasticity	.24	.32	.24	.30	.14	.18

NOTE: Table notes indicate significance of t ratios for coefficients (two-tailed test).
[1]> .01.
[2]> .05.
[3]> .10.
[4]t ratio greater than 1.0.

TABLE D-3
Alternative Specifications of Simple Expenditure Models for Owners of One-Family Units, Using Weighted and Unweighted Samples

	Linear		Semilog		Log-Log	
Variables	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
Race	−37.64[1]	−48.07[1]	.270[1]	−.361[1]	−.292[1]	−.350[1]
Income	7.48[1]	7.49[1]	.033[1]	.031[1]	.162[1]	.256[1]
Constant	107.50[1]	116.60[1]	9.38[1]	9.47[1]	8.24[1]	7.49[1]
R^2	.40	.40	.35	.40	.22	.32
Income elasticity	.39	.40	.28	.32	.16	.26

NOTE: Table notes indicate significance of *t* ratios for coefficients (two-tailed test).
[1]> .01.
[2]> .05.
[3]> .10.
[4]*t* ratio greater than 1.0.

Weighting has a similar impact on the income coefficients of the owner models. It increases the value of the income coefficient obtained for the log-log model considerably, i.e., from .162 to .256. The coefficient of the additive model increases, but by much less than in the additive renter equations, and the income coefficient in the semilog model is relatively unaffected by the use of weights.

A more complex set of results is obtained when the renter and owner equations are stratified by race (Tables D-4 and D-5). For black renters, weighting reduces the size of the income coefficients in all three specifications. The opposite result is obtained for white renters; the size of the income coefficient is always larger in the weighted regressions than in the unweighted regressions. Weighting affects the coefficients of income for black owners in the opposite way from coefficients for black renters; i.e., it increases the size of the income coefficient in all three specifications. Even so, with the exception of the semilog equations, the size of the income coefficients for black renters is smaller in nearly all cases than those obtained in either the weighted or unweighted forms of the comparable white equations.

Weighting has a mixed effect on the magnitude of the income coefficients for white owners. It generally has a smaller effect on the estimated income coefficients than was true of the other samples. Using weights decreases the size of the income coefficient in the additive and semilog specifications, but employing them in the log-log specification hardly changes the size of the income coefficient, which is .283 using uniform sample weights and .284 using population weights.

When both weighted and unweighted specifications are included, a total of 54 different income coefficients have been estimated. The 54 income elasticities implied by these estimates are summarized in Tables D-6 and D-7. The first of these tables presents 27 separate estimates for renters, and the second presents 27 separate estimates for owners. For the sample of all renter households and of white renter households, the income elasticities obtained from the weighted estimates are always larger than those obtained from unweighted estimates. For all but two cases, moreover, the elasticity calculated at the sample means from the linear models is larger than that calculated from the semilog models. The estimated elasticities from the semilog model are, in turn, larger than those obtained from the log-log model. In contrast, for black renters, the estimated income elasticities obtained from the unweighted simple models are larger than those computed from the weighted models. The largest income elasticity for black renters, .24, is obtained from the weighted estimate of the semilog specification of the simple model.

Comparing the results for white and black renters, the elasticities calculated from the linear and semilog specifications are larger for white

TABLE D-4
Alternative Specifications of Simple Expenditure Models for Black and White Renters, Using Weighted and Unweighted Samples

	Linear		Semilog		Log-Log	
Variables	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
			Blacks			
Income	2.61[1]	2.19[1]	.055[1]	.043[1]	.135[1]	.104[1]
Corrections						
No furniture	−4.04	−3.04	−.163[4]	.006	−.187[4]	−.003
No heat	−15.53[1]	−14.70[1]	−.304[1]	−.271[1]	−.337[1]	−.298[1]
No appliances	3.44	1.03	.222[2]	.008	.213[2]	−.007
No water	−1.73	−1.87	−.061[4]	−.031	.030	−.003
Constant	55.02[1]	60.26[1]	3.86[1]	3.983[1]	3.05[1]	3.36[1]
R^2	.346	.327	.28	.290	.22	.236
Income elasticity	.21	.17	.24	.19	.14	.10

Whites

Income	2.97[1]	3.94[1]	.040[1]	.044[1]	.134[1]	.202[1]
Corrections						
No furniture	8.69[4]	9.80[4]	.126[4]	.140	.185[3]	.232[2]
No heat	−6.90[3]	−3.66	−.156[1]	−.116[2]	−.164[1]	−.068[4]
No appliances	−32.80[1]	−28.52[1]	−.393[1]	−.339[1]	−.410[1]	−.432[1]
No water	6.93[4]	14.26[1]	.101[4]	.206[1]	.079[4]	.179[1]
Constant	74.02[1]	69.47[1]	4.21[1]	4.203[1]	3.30[1]	2.74
R^2	.355	.466	.33	.444	.27	.406
Income elasticity	.26	.33	.25	.33	.13	.20

NOTE: Table notes indicate significance of t ratios for coefficients (two-tailed test).
[1]> .01.
[2]> .05.
[3]> .10.
[4]t ratio greater than 1.0.

TABLE D-5
Alternative Specifications of Simple Expenditure Models for Black and White Owners, Using Weighted and Unweighted Samples

	Linear		Semilog		Log-Log	
Variables	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
			Blacks			
Income	3.45[1]	4.82[1]	.030[1]	.045[1]	.032	.080[2]
Constant	95.67[1]	85.15[1]	9.13[1]	9.02[1]	9.06[1]	8.62[1]
R^2	.112	.214	.13	.247	.02	.07
Income elasticity	.40	.41	.30	.32	.03	.08
			Whites			
Income	7.74[1]	7.54[1]	.033[1]	.031[1]	.283[1]	.284[1]
Constant	105.10[1]	116.10[1]	9.38[1]	9.47[1]	7.61[1]	7.23[1]
R^2	.372	.368	.29	.347	.20	.281
Income elasticity	.19	.26	.20	.28	.28	.28

NOTE: Table notes indicate significance of *t* ratios for coefficients (two-tailed test).
[1]> .01.
[2]> .05.

TABLE D-6
Summary of Estimated Income Elasticities for Renters

	Linear	Semilog	Log-Log
All			
Unweighted full model	.20	.18	.08
Unweighted simple model	.24	.24	.14
Weighted simple model	.32	.30	.18
Blacks			
Unweighted full model	.15	.19	.07
Unweighted simple model	.21	.24	.14
Weighted simple model	.17	.19	.10
Whites			
Unweighted full model	.25	.23	.09
Unweighted simple model	.26	.25	.13
Weighted simple model	.33	.33	.20
Black means and white coefficients			
Unweighted full model	.23	.15	–
Unweighted simple model	.24	.17	–
Weighted simple model	.31	.20	–

TABLE D-7
Summary of Estimated Income Elasticities for Owners

	Linear	Semilog	Log-Log
All			
Unweighted full model	.42	.26	.13
Unweighted simple model	.39	.28	.16
Weighted simple model	.40	.32	.26
Blacks			
Unweighted full model	.09	.07	–.00
Unweighted simple model	.19	.20	.03
Weighted simple model	.26	.28	.08
Whites			
Unweighted full model	.42	.28	.26
Unweighted simple model	.40	.30	.28
Weighted simple model	.41	.32	.28
Black means and white coefficients			
Unweighted full model	.45	.19	–
Unweighted simple model	.42	.21	–
Weighted simple model	.41	.19	–

than for black households. Moreover, the income elasticities for whites are higher than for blacks in two out of three of the log-log models.

These comparisons of black and white income elasticities may be misleading, since average sample characteristics are used in the elasticity calculations. The elasticity derived from the linear model, for example, is a linear function of the income/shelter ratio, while the elasticity derived from the semilog specification is a linear function of income. The differences in elasticities may be, therefore, less a reflection of behavioral differences between whites and blacks than of differences in the composition of the two groups. The elasticities presented in the last section of both tables—titled "black means and white coefficients"—computed by using the income coefficients of the several equations and black sample means, illuminate this issue.

This analysis indicates that if the coefficients of the white linear model are "correct," renters with income/shelter ratios equal to those of the mean black renter should have income elasticities in the range .23–.31. Instead we find that the estimated income elasticities of black renters in St. Louis fall in the range .15–.21. Similarly, use of the semilog specification of the white renter equation and the black sample means provides estimated income elasticities on the order of .15–.20, whereas the income elasticities calculated for black renters from the black semilog equation vary between .19 and .24.

In summary, for the weighted and unweighted versions of all three specifications, the observed income elasticities are, on the average, lower for black than for white renters. For the linear and log-log specifications, the estimated income elasticities are smaller for blacks than for whites with similar incomes and income/shelter ratios, but this result does not hold for the semilog comparison.

Similar results for the owners of single-family units are shown in Table D-7. In all three specifications for black owners, the weighted simple model yields larger elasticities than either the unweighted simple model or the unweighted full model. For all owners and for white owners, the income elasticities computed from the linear model are higher than those computed from either the semilog or log-log model. In contrast, elasticities for black owners computed from the semilog are slightly larger than those computed for the linear model. The calculated income elasticities are smaller for black than for white owners, for all nine comparisons. For the linear equations, the income elasticities for blacks computed using black means in white equations are slightly larger than those computed for whites. For the semilog specification, however, the income elasticities for blacks using white equations are considerably lower than those computed for whites.

In summary, Tables D-6 and D-7 indicate that income elasticities for owners are substantially higher than for renters, and that income elasticities are higher for white than for black households. The difference between the income elasticities calculated for whites and blacks is substantially larger for owners than for renters.

There is evidence, but it is not entirely consistent, that the income elasticity for black households is smaller than that of whites of similar socioeconomic characteristics. At the least, our analysis indicates a substantial difference in the intercepts and slopes of equations for white and black households.

Finally, the tables indicate that the magnitudes of the income elasticities are small when compared to results from aggregate analyses of the housing market. They are also slightly smaller than those obtained in Lee's microanalysis, but they are within the range of Straszheim's results.