Volume Title: Annals of Economic and Social Measurement, Volume 4, number 1

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm75-1

Publication Date: 1975

Chapter Title: Place to Place Rent Comparisons

Chapter Author: Robert Gillingham

Chapter URL: http://www.nber.org/chapters/c10222

Chapter pages in book: (p. 153 - 178)

Annals of Economic and Social Measurement, 4/1, 1975

PLACE TO PLACE RENT COMPARISONS

BY ROBERT GILLINGHAM*

This paper uses hedonic quality adjustment techniques to make interarea comparisons of rent levels. Hedonic equations were estimated for ten major cities using neighborhood quality characteristics as well as data on individual rental units. These regression results as well as the data on which they were estimated were used to construct several Laspeyres type, place to place rent comparisons. The effect on the rent comparisons of different coverage was investigated and hedonic based indexes were compared to indexes derived from data published by the Bureau of Labor Statistics.

1. INTRODUCTION[1]

Despite the fact that consumers must often make place to place comparisons of relative prices, there are very few unpublished data which measure geographic price differentials within the United States. The Consumer Price Index (CPI) provides price information for assessing changes in living costs within a large number of local areas over time, but it does not provide any information for assessing the living cost differences which exist among these areas. The purpose of this paper is to construct, as a first step toward providing such information, place to place rent comparisons among ten major U.S. cities.

This study will focus on the rent of central city, multiple unit dwellings. Data availability is an important criterion, but there are several other reasons for making this choice. First, rent expenditure in general, and multiple unit rent in particular, are important components of consumer expenditure, comprising approximately 5 and 4 percent of the consumer budget respectively. Second, rental prices are a good proxy for the cost of shelter for homeowners. To the extent that multiple unit and single family unit rents exhibit similar geographic differentials, the results from this study will also shed light on the much larger overall shelter component. Finally, there are several *a priori* reasons for believing that the shelter component of personal consumption is a prime source of cross-section variability in both price level and price change. Once in place, the capital stock necessary to provide housing services is immobile and difficult to alter. In addition, construction costs themselves exhibit significant geographic differentials and the lag time for new construction is relatively long. As a result, place to place differences in shelter costs cannot be diminished by competitive behavior as quickly or easily as price differences in other, less bulky, more easily transportable items.

---

* I wish to express my appreciation to Professor Robert Summers for his assistance throughout this project. In addition, I wish to thank Dale Heien, Robert Pollak, and Jack Triplett for helpful comments, and the staff of the Division of Consumer Prices and Price Indexes (BLS) for help in assembling and understanding the data used. I alone am responsible for any remaining errors. This paper does not represent an official position of the Bureau of Labor Statistics.

[1] This paper is a substantially abridged version of [2] which was, itself, a shortened and modified version of [1]. The empirical results which are only referred to in the present paper are presented in detail in both of the previous versions.

The only data now available which present estimates of geographic price differences are the City Worker's Family Budgets (CWFB) published by the Bureau of Labor Statistics [5], [6], [19]. These figures support the contention that regional variations in shelter costs are a major source of living cost differences. However, the CWFB estimates can only be taken as support and not verification because of the manner in which they are designed. The market baskets which are priced to obtain budget estimates are not defined to be representative of the expenditure of any particular group. Rather, they are chosen under a system of normative rules which attempt to define a particular standard of living for a tightly specified family group. The remainder of this paper will be an attempt to design a more general, broader purpose system of place to place rent comparisons.

A primary consideration in the construction of consumer price comparisons is, quite naturally, the definition of the consumption units to be priced. Because of the complexity of the housing services bundle, it is impossible to develop meaningful rent comparisons without an attempt to sort out at a more detailed level the characteristics of the service flow. The "new theory" of consumer demand forwarded by Kelvin Lancaster ([7], [8], [9]) provides a useful framework for analyzing the consumption process and, consequently, provides an interpretation for the hedonic quality adjustment technique which will be used to develop interarea rent comparisons.

Lancaster's "new theory" stems from three assumptions about consumer behavior. First, a good (e.g., shelter) does not give utility to a consumer, but rather embodies characteristics which are, in fact, the arguments of the consumer's preference ordering. Second, a good will, in general, embody more than one characteristic and share a given characteristic with other goods. Finally, the joint consumption of several goods may yield a different set of characteristics than will the independent consumption of the same goods. These assumptions yield a model of consumer behavior in which the preference ordering is defined in characteristics space, while the budget constraint is defined in goods space. The objective of the hedonic quality measurement technique is to establish a relationship between the observed market prices for goods and the implicit prices of the characteristics which the goods embody. Estimates of these implicit prices are obtained from a regression of goods prices on the associated quantities of embodied characteristics:

$$p_i = f(z_{i1}, \ldots, z_{im}) \qquad i = 1, \ldots, n. \tag{1.1}$$

Hedonic regressions of this type will be employed in this study to estimate the implicit prices of the characteristics embodied in the flow of services provided by rental units.[2] There are several ways in which these implicit prices could be used to develop place to place rent comparisons. Ideally, they could be used as data in the estimation of a set of characteristic demand functions obtained from a solution of Lancaster's consumer behavior model.[3] The results obtained from the estimation of the demand system could then be used to estimate an interarea

---

[2] The relationship between the "new theory" and hedonic analyses has been discussed in several places, e.g. Muellbauer [10], Rosen [14], and Triplett [15]. However, as will be pointed out below, the relationship is still unclear and warrants considerable additional research.

[3] A very interesting attempt in this direction has been made by A. Thomas King in [4].

cost of living subindex for rental unit shelter services.[4] This approach, however, presupposes an appropriate separability of the preference ordering. Furthermore, it requires that the Lancaster model, including its consumption technology aspects, be specified more completely and in a form which admits to solution. Unfortunately, the theoretical problems inherent in this approach have not been solved, nor even clearly defined. Consequently, in this study the implicit characteristic prices will be used in the construction of Laspeyres-type place to place rent comparisons.

There are several reasons for following this approach. First, a Laspeyres place to place rent index is an important component of a complete interarea Laspeyres index which can be used to define an upper bound on the cost of living differential between two areas. Second, if the separability conditions which would be required to estimate the characteristic demand functions described above do in fact exist, the Laspeyres rent index will define an upper bound on the cost of living subindex corresponding to rental unit shelter services.[5] Finally, one of the objectives of this study is to define a methodology for making place to place rent comparisons among a substantial number of areas on a continuing basis. The Laspeyres framework provides a conceptually straightforward and operationally feasible means for achieving this goal.

2. ESTIMATION OF THE HEDONIC RENT REGRESSIONS

The quality of the shelter services provided by a rental unit is determined by four types of characteristics which can be cross-classified in the following manner:

	Physical Characteristics	Household Characteristics
Individual Unit		
Neighborhood		

The hedonic technique was used to estimate the effect on rent (i.e. the implicit price) of these four types of characteristics. The major objective of the estimation process was to assemble data on a sufficient number of quality variables to avoid a major impact of undefined and unstable proxy relationships in the analysis. It is virtually impossible to eliminate proxy relationships, but an attempt was made not only to minimize their incidence but to formulate them in such a way that they are unlikely to vary substantially from place to place.

The data with which the hedonic regressions were estimated were drawn from two major sources: (1) Microdata on individual housing units were drawn from the 1960–61 Comprehensive Housing Unit Survey conducted by the Bureau of Labor Statistics. (2) Neighborhood characteristics data were drawn from the

---

[4] The term subindex is used to refer to an index for a particular category of goods. Cost of living subindexes are derived by Robert Pollak in [11].

[5] A precise statement of the relationship between Laspeyres and cost of living subindexes is given in Pollak [11].

1960 decennial Census. Preliminary regression experiments were conducted to indicate data deficiencies and focus on a useful set of quality variables. The final list of variables, drawn from both sources, can be found in Appendix I. These variables were used to estimate hedonic equations for ten cities: Chicago, Los Angeles, Detroit, Boston, Pittsburgh, Cleveland, Washington, Baltimore, St. Louis, and San Francisco. Final parameter estimates for these equations are given in Appendix II. Although the estimation process will not be described in detail in this paper, several aspects of the process do warrant individual discussion.

In estimating the hedonic equations, considerable emphasis was placed on experimentation with alternative functional forms.[6] Although, there is nothing in hedonic "theory" which dictates the functional form of the regression equation, most previous analyses have used either a linear, semilog linear, or double log linear function.[7] In most cases the choice among these forms has been made on either *a priori* grounds or using explanatory power as the base criterion. In this study, the choice of functional form was made not only on these bases, but also on the degree to which the ordinary least squares assumptions of linearity and error variance homoscedasticity were satisfied. The choice was simplified by the fact that the predominance of dummy variables in the hedonic equations made the semilog and double log functional forms virtually indistinguishable. Because of this, and because there was no objective method for transforming the continuous variables of the equation into the strictly positive domain in order to estimate the double log form, the choice was limited to the linear and semilog formulations.

On *a priori* grounds, the semilog functional form has substantial appeal relative to the linear form. When our inability to measure the varying quality of the quality characteristics themselves is considered, use of the semilog formulation can be interpreted as an assumption that the quality, and thus the implicit cost, of these attributes is related to the rent level. For example, in the semilog model, the cost of an additional bathroom is not constant as the linear equation would imply, but rather varies proportionately with the rent level of the apartment.

Comparison of the statistical properties of the linear and semilogarithmic functional forms indicates that neither form exhibits clearcut superiority in either explanatory power or in the degree to which the ordinary least squares linearity assumption is satisfied.[8] However, the semilog transformation does appear to correct for a common form of error variance heteroscedasticity exhibited by the linear form in which the standard error is correlated with the conditional expectation of the dependent variable.[9] For this reason, and because of its *a priori* appeal, the semilog functional form was used to estimate the hedonic regressions used in this study.

Although we will not attempt a detailed discussion of the parameter estimates given in Appendix II, two comments on the estimated coefficients are in order. First, the individual coefficients have, for the most part, *a priori* correct signs and

---

[6] Parameter estimates for the linear functional form as well as hedonic equations estimated without neighborhood characteristics variables are given in [1] and [2].

[7] Throughout this study natural logarithms are used.

[8] Linearity of the estimated regressions was tested using the Durbin–Watson test suggested by Prais and Houthakker [13].

[9] A modification of a test suggested by Glejser [3] was used to test for the existence of this form of heteroscedasticity.

are of reasonable magnitude. There are, however, exceptions to this rule; for example, the coefficient on the garage availability variable in the Washington regression is negative and greater than its standard error. These exceptions indicate an inability to obtain data on enough quality characteristics to completely avoid all proxy relationships. Second, although the degree of multicollinearity evidenced in the rental unit data was less, in general, than that which has been common in previous hedonic studies, it did cause a problem with the treatment of certain groups of variables, most notably those representing the inclusion of household durables in rent. It was the high degree of collinearity among these variables which necessitated the construction of the particular household durable variables defined in Appendix I.

Neighborhood variables were included in the hedonic regression in order to avoid an incomplete specification in which included individual unit variables might act as proxies for excluded neighborhood information. Since neighborhood information is difficult to obtain for non-Census years, it is important to point out the effect inclusion of these variables had on the estimated regressions. The effect on most coefficient estimates was moderate, indicating perhaps that the excluded variable proxy problem is less severe than might be expected. However, in several cases, inclusion of neighborhood variables resulted in a substantial change in parameter estimates, almost always in an *a priori* reasonal direction.

The inclusion of neighborhood characteristics had its most significant effect on the estimated rent differentials for nonwhite households, increasing these estimates substantially. On average, the nonwhite occupants in the CHUS samples lived in neighborhoods which, according to the neighborhood variables included here, were of lower quality. When the neighborhood characteristics are included in the analysis the differential quality levels are accounted for, and the estimated equations indicate that, in most cases, nonwhite occupants pay more for comparable housing than white occupants. Accordingly, rent differentials estimated from equations using the individual housing unit data alone can be extremely misleading.

The inclusion of neighborhood characteristics increased the average R^2 for the ten cities by 0.06, thereby reducing the average amount of unexplained variance by approximately 15 percent. Perhaps more interesting is the fact that the average Durbin–Watson statistic, computed when the observations are ordered by block, is increased from 1.50 to 1.62.[10] When neighborhood variables are not included in the analysis it is plausible to assume that estimation errors for rents within the same block are not independent. The increase in the Durbin–Watson statistic indicates that at least a portion of the neighborhood variation leading to this dependence has been accounted for.

3. PLACE TO PLACE RENT COMPARISONS

The parameter estimates given in Appendix II provide a basis for pricing out a wide range of rental unit "specifications." The number of rent indexes which could be produced in this fashion is, of course, virtually unlimited. In

---

[10] The Durbin–Watson statistics given in Appendix II are computed on a block ordering.

this section we will concentrate on Laspeyres-type indexes which summarize in several ways the rent differences among the cities under study. Two types of indexes were computed. The first set estimates a matrix of place to place rent indexes for the types of units which were rented in each of the ten cities. These indexes demonstrate, among other things, the sensitivity of place to place rent indexes to differences among the types of rental units for which they are computed. The second set of indexes provides a summary index estimating place to place rent differences constructed for the total sample of rental units specifications in the ten cities studied and, in addition, a similar index which focuses on the set of five room units meeting the specifications of the City Worker Family Budgets mentioned in Section 1. The purpose of the summary index is to provide a single overall indicator of the place to place variability of multi-unit apartment rents. The purpose of the more specific index is to provide a hedonic based index which is similar in coverage to a rent index derived from the 1959 CWFB, with which it will be compared.

The Laspeyres index in which we are interested is the ratio of the costs of renting a specified set of rental unit characteristics under two price regimes. Following Pollak [12], we will define these regimes as the reference and comparison situations, and the particular rental unit specification which we want to cost out is one which is actually rented in the reference situation. For our first set of indexes, the reference and comparison situations will each be defined as one of the ten cities. Let x^s denote a vector of characteristics which describes a multi-unit apartment in city s. If we let $R^t(x^s)$ and $R^s(x^s)$ denote rents in cities t and s for the rental unit described by x^s, then a Laspeyres rent index comparing city t to city s can be written as:

$$\Lambda_{st}(x^s) = R^t(x^s)/R^s(x^s). \tag{3.1}$$

A separate Laspeyres index can be computed for each of the multi-unit apartments in city s. These indexes will vary for two reasons. First, the vector of rental-unit characteristics varies over renter households, and, second, rent for a particular vector of characteristics is subject to variation in both the reference and comparison cities. However, the parameters of the distribution of the logarithm of Λ_{st} can be derived quite easily, and these parameters can be used to describe the distribution of Λ_{st}.

Using the hedonic equation for each of the cities, we can write

$$\ln R^t(x^s) = \sum_{j=1}^{m} \alpha_j^t x_j^s + u^t \tag{3.2}$$

and

$$\ln R^s(x^s) = \sum_{j=1}^{m} \alpha_j^s x_j^s + u^s \tag{3.3}$$

where m is the number of characteristics and α^t, α^s, u^t, and u^s are coefficient vectors and disturbance terms from the hedonic equations for cities t and s. Taking the logarithm of equation (3.1) and substituting (3.2) and (3.3) into the result yields

$$\lambda_{st}(x^s) = \ln \Lambda_{st}(x^s) = \sum_{j=1}^{m} (\alpha_j^t - \alpha_j^s)x^s + u^t - u^s. \tag{3.4}$$

The expected value of λ_{st} can be written as

$$(3.5) \qquad E(\lambda_{st}) = \sum_{j=1}^{m} (\alpha_j^t - \alpha_j^s)E(x_j^s).$$

Assuming that x^s is a nonstochastic vector, and that u^s and u^t are independently distributed, the variance of λ_{st} is given by

$$(3.6) \quad \operatorname{var}(\lambda_{st}) = \sum_{j=1}^{m} \sum_{k=1}^{m} (\alpha_j^t - \alpha_j^s)(\alpha_k^t - \alpha_k^s) \operatorname{cov}(x_j^s, x_k^s) + \operatorname{var}(u^t) + \operatorname{var}(u^s).$$

The antilog of equation (3.5) is the geometric mean of λ_{st}.[11]

To compute estimates of the expected value and variance of λ_{st}, we used the same sample of multi-unit apartment specifications with which the hedonic equations were estimated. The expected value of λ_{st} was estimated by

$$(3.7) \qquad \bar{\lambda}_{st} = \sum_{j=1}^{m} (\hat{\alpha}_j^t - \hat{\alpha}_j^s)\bar{x}_j^s$$

where $\bar{x}_j^s$ is the sample mean of x_j^s and α_j^t and $\hat{\alpha}_j^s$ are the least squares estimates of α_j^t and α_j^s. The variance of λ_{st} was estimated by

$$(3.8) \qquad \sigma^2_{\lambda_{st}} = \sum_{j=1}^{m} \sum_{k=1}^{m} (\hat{\alpha}_j^t - \hat{\alpha}_j^s)(\hat{\alpha}_k^s - \hat{\alpha}_k^s)\hat{\sigma}^2_{x_j^s, x_k^s} + \hat{\sigma}^2_{u^t} + \hat{\sigma}^2_{u^s}$$

where $\hat{\sigma}^2_{x_j^s, x_k^s}$ is the estimated cov (x_j^s, x_k^s) computed from the regression data and $\hat{\sigma}^2_{u^t}$ and $\hat{\sigma}^2_{u^s}$ are the estimated error variances from the hedonic regressions. The estimated geometric mean of Λ_{st}, $\bar{\Lambda}^g_{s_t}$, was computed by exponentitating equation (3.7). The estimated variance of $\bar{\lambda}_{st}$ is

$$(3.9) \qquad \hat{\sigma}^2_{\bar{\lambda}_{st}} = \hat{\sigma}^2_{\lambda_{st}}/n$$

where n is the number of rental units in the s-th city sample. Limits of the 95 percent confidence intervals for $\bar{\Lambda}^g_{st}$ were estimated by exp $\{(\bar{\lambda}_{st} \pm 2\hat{\sigma}^2_{\bar{\lambda}_{st}})\}$.

In actuality, the estimated geometric means were computed for the subset of renter unit specifications in the reference city for which a cost could be estimated in both the comparison and reference cities. When the number of units in either city sample possessing a particular characteristic was so small that a coefficient for this attribute could not be estimated, rental unit specifications from the reference city sample which included this attribute were not used in computing either $\bar{x}^s$ or the estimated covariance matrix of x^s, and were thus, in effect, "linked out" of the rent comparison. For example, since the Boston rental unit sample had only two units with more than one bathroom, a reliable coefficient for this attribute could not be estimated in the Boston regressions. Consequently, when Boston rent levels were compared to those of the other cities in which a coefficient for this attribute could be estimated, specifications which included more than one bathroom were not included in the set of reference city specifications used in computing

---

[11] Under the assumption that λ_{st} is normally distributed, the geometric mean of Λ_{st} would be equal to the median of Λ_{st}.

the index. In most cases, a rent for nearly all reference city specifications could be estimated in the comparison city.

Table 3.1 gives numerical estimates of $\bar{\Lambda}^g_{st}$. The numbers in parentheses define 95 percent confidence intervals for $\bar{\Lambda}^g_{st}$. The last row in Table 3.1 gives the average of $\bar{\Lambda}^g_{st}$ over all t, for each s. The figure in parentheses in this row is the standard deviation of the ten index levels in each column. Each column of Table 3.1 gives the estimated geometric mean index in each of ten cities for the set of rental unit specifications purchased by a different group of renter families, i.e., the renter population of the reference city specified in the column heading. The degree to which these geometric mean indexes change in response to a change in the reference city is striking. First, inspection of the standard deviation of index levels given in the last row of Table 3.1 indicates that the degree of index level variation exhibited by each of the indexes varies substantially as the reference city is changed. Second, each of the indexes in Table 3.1 implies a different rank ordering for rent levels in the ten cities. The rank order correlation between each pair of indexes in Table 3.1 was computed and the values of these correlation coefficients do indicate that the rank orderings of the indexes in each of the ten columns of Table 3.1 are, in general, positively related to one another.[12] However, the make-up of the stock of apartments differs sufficiently from city to city to yield an average rank order correlation coefficient between indexes of only 0.496.

The index differences just described are what is gained by partitioning the population of apartment renters by city of residency. On average, an apartment renter would get a better estimate of a Laspeyres index which was specific to his situation by looking at an index in which his city was the reference city than he would if he could only refer to a single aggregate place to place rent index. Despite these gains, however, indexes defined for particular households or particular rental units specifications are still subject to wide variation around the geometric mean indexes given in Table 3.1. An estimated "two-sigma" confidence interval was computed for each Λ_{st}, where the limits of the interval are defined as exp $(\{\bar{\lambda}_{st} \pm 2\sigma_{\lambda_{st}})\}$. On average, these intervals run from 53.09 to 196.76! The size of the estimated variance of λ_{st} and, thus, the width of these confidence intervals results to a substantial degree from within city variance of rents for identical units estimated from the hedonic regressions.

Thus, even though the geometric mean indexes presented in Table 3.1 can be accurately estimated, rent indexes for specific renter families or specific rental unit specifications in each of the reference cities are subject to wide variation. The indexes given in Table 3.1 are a convenient and reasonable first step toward constructing disaggregated place to place rent indexes which can be more useful to the reference groups for which they are defined. The differences among these

---

[12] The rank order correlation coefficient between the index for reference city j and the index for reference city k was computed by

$$1 - \sum_{t=1}^{10} d_t^2/(n^2 - n)$$

where n was equal to 10 and d_t was the difference between the ranks of $\bar{\Lambda}^9_{jt}$ and $\bar{\Lambda}^9_{kt}$. For a one-tailed test of the null hypotheses that a rank order correlation coefficient in Table 3.1 is equal to zero (H_a: $r > 0$), the critical values for probabilities of 0.25, 0.10, and 0.05 are 0.248, 0.455, and 0.564 respectively [20, p. 579]. The full matrix of rank order correlation coefficients is given in [1] and [2].

indexes discussed above indicate that there is, in fact, a substantial information gain from disaggregation. However, the large within reference group variation which remains in these indexes indicates that future research should concentrate on developing a method of partitioning (in this case) apartment renters in a manner which yields indexes with smaller variances, thus yielding greater information gains from disaggregation for which the additional index construction costs which go along with disaggregation can be more easily justified.

In addition to the matrix of Laspeyres indexes given in Table 3.1, we are also interested in constructing summary indexes which provide a single vector of place to place rent comparisons using a sample of specifications from all ten cities to define a composite reference situation. Two such indexes were constructed. The first type is designed to provide an index equally representative of all multi-unit apartment renters in the ten cities under study. The second index is designed to perform the same function, except that the coverage is limited to five room apartments of a type specified by the BLS City Worker's Family Budget. Each of these indexes will be compared with a rent index derived from the 1959 CWFB, as well as an index of CHUS sample mean rents.

In order to construct the broad coverage index, all of the rental unit specifications in each of the individual city samples which, with two exceptions, could be priced in all of the ten cities were combined into a composite specification sample. Under this criterion, units with the following attributes were excluded from the ten city composite sample:[13]

(1) units with more than one bath (BATHS)
(2) furnished units with only a refrigerator *or* a freezer (HHDUR*)
(3) units without hot and cold running water (NOWATER)
(4) units without installed heat (NONINST).

The two exceptions were rental unit specifications which included either central air conditioning or an elevator. These attributes were treated differently because in each case only one city sample had too few units with the attribute to estimate a price for it, and the estimated hedonic equations indicated that in each case the attribute was an important determinant of rent in the majority of the other nine cities. In the index formulae which follow, the coefficients on the elevator building characteristic in St. Louis and central air conditioning in San Francisco were assumed to be equal to the weighted average of the coefficients for each of these characteristics in the other nine cities.

A similar composite specification sample was constructed for computation of the CWFB type index specified above. This sample was defined as the subset of the composite sample defined above meeting the following additional criteria:[14]

(1) five room unit
(2) one full bath, and
(3) sound condition.

---

[13] These four characteristics, BATHS, HHDUR*, NOWATER, NONINST, were included in an average of 2.3, 1.3, 3.1, and 5.4 percent, respectively, of the specification in each of the cities.

[14] The complete list of CWFB criteria also included "furnished," and "with complete kitchen facilities." These two criteria were not enforced for two reasons. First, the resulting sample sizes would have been extremely small. Second, these characteristics do not imply a particular type of structure but rather a particular type of rental market, and it was therefore felt that inclusion of units which did not fulfill these criteria did not contradict the basic objectives of the CWFB definition.

TABLE 3.1

ESTIMATED GEOMETRIC MEAN RENT INDEXES FOR 1960
(95 percent confidence intervals in parentheses)

Reference City[1] Comparison City[1]	Chicago	Los Angeles	Detroit	Boston	Pittsburgh
1. Chicago	100.0 —	97.69 (93.77, 101.78)	125.28 (122.73, 127.88)	102.65 (98.54, 106.24)	100.08 (96.38, 103.92)
2. Los Angeles	107.14 (104.89, 109.43)	100.0 —	110.63 (108.54, 112.77)	109.15 (104.44, 114.07)	109.37 (105.73, 113.13)
3. Detroit	92.01 (90.20, 93.86)	78.56 (75.49, 81.76)	100.0 —	92.07 (88.77, 95.49)	98.62 (94.97, 102.43)
4. Boston	102.83 (100.56, 105.15)	99.34 (95.25, 103.61)	108.41 (106.14, 110.72)	100.0 —	100.57 (95.21, 106.23)
5. Pittsburgh	95.78 (93.76, 97.84)	102.97 (98.68, 107.44)	118.81 (116.18, 121.49)	110.16 (105.55, 114.97)	100.0 —
6. Cleveland	91.71 (89.97, 93.48)	99.44 (95.61, 103.42)	115.73 (113.54, 117.96)	96.00 (92.22, 99.94)	104.78 (101.06, 108.64)
7. Washington	94.70 (92.68, 96.76)	108.33 (103.67, 113.21)	116.27 (113.91, 118.69)	102.26 (97.51, 107.24)	113.19 (109.06, 117.49)
8. Baltimore	91.20 (89.34, 93.11)	87.69 (84.09, 91.45)	103.64 (101.49, 105.83	93.22 (89.46, 97.14)	98.80 (95.20, 102.53)
9. St. Louis	95.54 (93.55, 97.56)	92.42 (88.91, 96.07)	112.23 (109.84, 114.67)	97.62 (93.28, 102.16)	99.33 (95.76, 103.03)
10. San Francisco	93.15 (90.58, 95.79)	96.39 (92.19, 100.78)	104.55 (102.14, 107.02)	97.25 (92.80, 101.90)	92.56 (88.45, 96.86)
Average (Std. Dev. of Levels)	96.41 (4.93)	96.28 (7.92)	111.56 (7.33)	100.04 (5.81)	101.73 (5.61)

TABLE 3.1 (continued)

Reference City[1] Comparison City[1]	Cleveland	Washington	Baltimore	St. Louis	San Francisco
1. Chicago	106.47 (103.96, 109.05)	120.27 (117.21, 123.42)	111.75 (109.01, 114.56)	95.81 (93.21, 98.49)	98.82 (94.80, 103.01)
2. Los Angeles	103.62 (101.28, 106.00)	108.30 (105.78, 110.88)	101.70 (99.39, 104, 07)	110.58 (107.73, 113.51)	106.49 (103.04, 110.05)
3. Detroit	90.23 (88.16, 92.34)	103.24 (100.73, 105.81)	100.24 (97.80, 102.74)	92.77 (90.39, 95.21)	88.94 (85.64, 92.37)
4. Boston	97.21 (94.53, 99.97)	129.12 (124.89, 133.49)	108.74 (105.25, 112.34)	88.06 (85.62, 90.56)	104.22 (99.70, 108.94)
5. Pittsburgh	97.90 (95.38, 100.48)	115.93 (113.11, 118.82)	108.62 (105.90, 111.41)	98.65 (96.15, 101.20)	96.45 (92.58, 100.48)
6. Cleveland	100.0 —	115.90 (112.91, 118.97)	111.24 (108.49, 114.07)	102.09 (99.40, 104.85)	93.44 (89.87, 97.15)
7. Washington	101.70 (99.31, 104.15)	100.0 —	110.48 (107.88, 113.14)	118.62 (115.34, 121.99)	101.98 (97.87, 106.26)
8. Baltimore	94.07 (91.81, 96.38)	101.97 (99.54, 104.46)	100.0 —	96.73 (94.35, 99.18)	92.57 (88.98, 96.31)
9. St. Louis	100.32 (97.84, 102.87)	103.95 (101.16, 106.81)	102.97 (100.48, 105.53)	100.0 —	95.48 (91.73, 99.38)
10. San Francisco	92.83 (89.88, 95.87)	102.57 (99.27, 105.98)	95.65 (92.64, 98.75)	92.13 (89.10, 95.27)	100.0 —
Average (Std. Dev. of Levels)	98.44 (4.77)	110.13 (9.20)	105.14 (5.40)	99.54 (8.65)	97.84 (5.22)

[1] The reference city (given in the column heading) is the city from which the sample of rental unit specifications is drawn, i.e., the reference situation in the Laspeyres framework. The comparison city is equivalent to the comparison situation in the Laspeyres framework. Thus each column contains an estimated geometric mean Laspeyres index comparing rents in the cities specified in each row to rents in the city specified in the column head.

These two specification samples were priced in each of the ten cities, and the estimated rent in each of the ten cities for each of the specifications was compared to the weighted ten city average rent of that specification. In other words, the weighted ten city average rent was used to represent the "average" reference situation. Weights were computed as the ratio of the number of multi-unit rental dwellings in a city to the total number of such dwellings in the ten cities. The following weights were derived from 1960 Census data [16]:

Chicago	0.353	Cleveland	0.059
Los Angeles	0.154	Washington	0.070
Detroit	0.077	Baltimore	0.033
Boston	0.076	St. Louis	0.065
Pittsburgh	0.033	San Francisco	0.082

An estimated geometric mean index was computed over all the specifications in both the overall composite specification sample and the CWFB-type specification sample. In the computation process the specifications themselves were weighted according to the city sample from which they were originally drawn. The estimated geometric mean index comparing city t to the ten city weighted average can be written as

$$\overline{\Lambda}^{g}_{*t} = \exp(\bar{\lambda}_{*t}) = \exp\left\{ \sum_{j=1}^{m} (\hat{\alpha}^{t}_{j} - \hat{\alpha}^{*}_{j}) \sum_{i=1}^{10} w_i \bar{x}^{i}_{j} \right\} \tag{3.10}$$

where w_i is the weight of the i-th city, $\hat{\alpha}^{*}_{j} = \sum_{i=1}^{10} w_i \hat{\alpha}^{i}_{j}$, and $\bar{x}^{i}_{j}$ is the mean of the j-th characteristic for those observations in the composite sample from city i. Of course the mean vector of characteristics, $\bar{x}^{i}$, differs according to whether the overall index or the CWFB-type index is being computed.[15] Assuming that the error structures of the hedonic equations are independent, the limits of a 95 percent confidence interval for $\overline{\Lambda}^{g}_{*t}$ can be estimated by $\exp(\bar{\lambda}_{*t} \pm 2\hat{\sigma}^{2}_{\bar{\lambda}_{*t}})$, where

$$\hat{\sigma}^{2}_{\bar{\lambda}_{*t}} = \sum_{i=1}^{10} w_i^2 (1/n_i) \left[\sum_{j=1}^{m} \sum_{k=1}^{m} (\hat{\alpha}^{t}_{j} - \hat{\alpha}^{*}_{j})(\hat{\alpha}^{t}_{k} - \hat{\alpha}^{*}_{k}) \hat{\sigma}^{2t}_{x_j, x^t_k} + \hat{\sigma}^{2}_{u^t} + \hat{\sigma}^{2}_{u^*} \right] \tag{3.11}$$

and $\hat{\sigma}^{2}_{u^*} = \sum_{i=1}^{10} w_i^2 \hat{\sigma}^{2}_{u^i}$.

Table 3.2 presents indexes computed from both the overall and CWFB-type composite specification samples. Limits of the 95 percent confidence intervals are given in parentheses for these two indexes. In addition to the two regression based indexes, Table 3.2 also contains an index derived from the 1959 interim CWFB, and an index of CHUS sample mean rent levels for the ten cities. It is important to reiterate that the regression based CWFB-type index is not strictly comparable to the official BLS CWFB. The two differ slightly in their definition of acceptable units (cf. footnote 14), and the official CWFB index covers units which are not in the central city as well as single family units. The mean and standard

---

[15] The same weights were used for the construction of both the overall and the CWFB-type indexes. Because data are not available on the relative number of five room multi-unit apartments in each of the cities, it was implicitly assumed that the proportion of such units in each city was the same.

TABLE 3.2

ESTIMATED PLACE TO PLACE RENT INDEXES FOR 1960

(95 percent confidence in intervals in parentheses)

City / Type Index	Full Specification Sample	CWFB-Type Sample	1959 BLS CWFB	CHUS Sample Average
1. Chicago	102.54 (101.69, 103.40)	104.07 (101.24, 106.98)	111.95	109.71
2. Los Angeles	104.63 (103.83, 105.44)	103.93 (101.49, 106.42)	95.15	93.63
3. Detroit	90.02 (89.31, 90.74)	87.89 (85.58, 90.26)	84.00	89.86
4. Boston	101.94 (100.96, 102.93)	99.94 (96.94, 103.04)	100.15	99.30
5. Pittsburgh	100.03 (99.04, 101.03)	92.45 (89.58, 95.41)	81.74	83.67
6. Cleveland	97.69 (96.89, 98.50)	96.04 (93.47, 98.68)	96.20	100.99
7. Washington	100.10 (99.22, 100.99)	95.89 (93.28, 98.57)	99.02	107.06
8. Baltimore	91.57 (90.80, 92.35)	92.37 (89.77, 95.03)	81.09	93.30
9. St. Louis	96.76 (95.93, 97.59)	95.88 (93.13, 98.71)	104.84	76.39
10. San Francisco	96.47 (95.42, 97.53)	103.88 (100.37, 107.52)	87.15	101.54
Weighted Average	100.00	100.00	100.00	100.00
(Wtd. Std. Dev. of Levels)	(4.22)	(5.26)	(10.58)	(9.87)
Unweighted Average	98.18	97.23	94.13	95.55
(Unwtd. Std. Dev. of Levels)	(4.33)	(5.38)	(9.85)	(9.81)

deviation, both weighted and unweighted, for each of the two indexes in Table 3.2 are given in the last two rows of the table.

The standard deviations given in the last rows of Table 3.2 indicate that estimation of the place to place variability of rent levels is sensitive to both the coverage of the index and the computational method. Both of the regression based indexes exhibit substantially less place to place variation in rent levels than the official CWFB index. Furthermore, the hedonic index with the broader coverage exhibits less variation than the CWFB-type index. The differences between the place to place variability of the official CWFB index and the place to place variability of the hedonic indexes are indeed substantial. Comparison of the place to place variability of the broad coverage hedonic index with the place to place variability of the index of CHUS sample mean rents indicates that the quality corrected rent levels exhibit considerably less variability and, thus, a substantial part of the variability of the index of sample means is due to variation in average quality level. The rank order correlation between the full sample hedonic index and the 1959 CWFB index is 0.552, while the correlation between the CWFB-type hedonic index and the official CWFB index is 0.588. It is unlikely that the low correlation between the BLS CWFB index and the CWFB-type index can be caused by differences in coverage alone, since the rank order correlation between the two regression-based indexes is 0.745.

4. SUMMARY AND CONCLUSIONS

The purpose of this study was to develop a method for making place to place rent comparisons among ten large U.S. cities using hedonic techniques. The comparisons were developed using the hedonic quality adjustment technique which was put, at the outset, within the framework of the characteristics approach to the analysis of consumer behavior introduced in Lancaster's "new theory."

The major objective in defining the hedonic rent function was to specify the relationship as completely as possible in order to avoid undefined proxy relationships. To do this, data on individual unit characteristics from the 1960–61 BLS Comprehensive Housing Unit Survey were combined with neighborhood characteristics data drawn from the 1960 Census. While the characteristics of housing are so complex that virtually no specification is complete, the CHUS and Census data, used in a single equation, provide a reasonable approximation to the rent determination process, and are the most complete data available for a large number of cities.

The estimated hedonic equations provide a basis for computing a network of Laspeyres place to place rent indexes. In theory, the Laspeyres (or Paasche) index for a given household can differ from that of any other household. In Section 3, estimated geometric mean Laspeyres indexes were computed for ten different sets of rental units—those occupied by the renter populations of each of the ten cities. The results, presented in Table 3.1, indicate that there is substantial variation among place to place rent indexes constructed for different reference groups, and that specification of the group to be represented by the index is a crucial aspect of index design. However, it was also shown that the construction of indexes under a partitioning of rental units by city yields indexes with a high degree of within group variation. Future research should be aimed at developing disaggregation methods which will yield indexes with low variances so that households within the coverage of an index will have a measure which is not only representative of them in the expectational sense, but also a close approximation to a measure which is designed specifically for them.

TABLE 4.1

ESTIMATED PLACE TO PLACE RENT INDEXES FOR 1967

City	Full Specification Sample	CWFB-Type Sample	Spring 1967 CWFB
1. Chicago	99.75	101.14	106.96
2. Los Angeles	105.82	105.00	101.20
3. Detroit	86.49	84.36	80.14
4. Boston	109.01	106.76	99.42
5. Pittsburgh	98.04	90.52	78.39
6. Cleveland	91.17	89.53	89.67
7. Washington	103.90	99.44	96.58
8. Baltimore	90.43	91.12	96.64
9. St. Louis	92.96	92.02	90.10
10. San Francisco	107.50	115.64	115.08
Weighted Average	100.00	100.00	100.00
(Wtd. Std. Dev. of Levels)	(6.48)	(7.95)	(9.75)
Unweighted Average	98.51	97.82	95.42
(Unwtd. Std. Dev. of Levels)	(7.55)	(9.13)	(10.76)

In addition to the "city" indexes described above, two types of summary rent indexes were constructed. The first type was designed to estimate a geometric mean rent index based on a composite specification sample drawn from all ten cities. The second type was designed to estimate a similar measure except that only those rental units (approximately) meeting City Worker's Family Budget specifications were included in the coverage. These indexes, presented in Table 3.2, demonstrated the effect of different estimation methods and different coverage on the place to place rent index. The full coverage hedonic index exhibited substantially less variation than the official CWFB index; in fact, the variance of the full coverage hedonic index was of the same order of magnitude as that exhibited by the non-shelter components of the CWFB.

The complete set of individual unit and neighborhood characteristics used to estimate the hedonic equations specified in this paper have not been available since the CHUS was collected. However, data collection requirements now under development for the Consumer Price Index, as well as the Annual Housing Survey of the Department of Housing and Urban Development, should provide the necessary data bases in the future. A rough and ready approximation of the changes which have occurred in the place to place indexes because of differential rent movements over time in the ten cities can be obtained by updating the indexes with the CPI rent component for each of the cities.[16] Table 4.1 presents an updating to 1967 of the two hedonic indexes presented in Table 3.2, as well as the spring 1967 CWFB index for the ten cities.

Office of Prices and Living Conditions
Bureau of Labor Statistics

APPENDIX I. GLOSSARY OF VARIABLES

I. CHUS VARIBLES (unless continuous, as indicated by (*) after name, definition states condition under which variable equals one):

1. 50–50: Unit built in years 1950 to 1960.
2. 40–50: Unit built in years 1940 to 1950.
2. 30–39: Unit built in years 1930 to 1939.
4. PRE1920: Unit built before 1920.
2. POST1940: Unit built in years 1940 to 1960 (used only for Boston and Pittsburgh regressions).
6. 2RM: Two room apartment.
7. 3RM: Three room apartment.
8. 4RM: Four room apartment.
9. 5RM: Five room apartment.
10. MT5RM: Unit has more than five rooms.
11. BATHS: Unit has more than one full bathroom.
12. NOBATH: Unit has less than one full bathroom.
13. HHDUR: Unit has furniture, refrigerator, *and* stove included in rent.
14. HHDUR*: Unit has furniture and a refrigerator *or* a stove included in rent.
15. MA: For unfurnished units, major appliances included in rent; 0 = none, 1 = refrigerator *or* stove, 2 = refrigerator *and* stove used with variable (15).

---

[16] The reasons why this is only a rough and ready technique are numerous. Among them are the fact that the CPI rent indexes cover all structure types and the total SMSA, and certainly cannot be simultaneously representative of both the overall specification sample and the CWFB type specification type.

16. COND: Unit in deteriorating or dilapidated condition.
17. NOWATER: Unit does not have cold *and* hot running water facilities in structure.
18. NONCENT: Unit has noncentral heat.
19. NONINST: Unit has no installed heat.
20. GARAGE: Unit has garage included in rent.
21. GARAVAIL: Unit has garage available to tenant but not included in rent.
22. CAC: Unit is in centrally air conditioned structure.
23. CAC*: Central air conditioning included in rent (used only for Washington regression).
24. ELEV: Unit is in elevator building.
25. PER/RM(#): Number of persons per room.
26. RACE: Unit has nonwhite head of household.

II. CENSUS VARIABLES (all continuous):

27. PNWU: Proportion of units on block occupied by nonwhite head of household.
28. PLACK: Proportion of units on block lacking (at least some) plumbing facilities.
29. PCRWD: Proportion of units on block which have an occupant density of greater than one person per room.
30. PSFU: Proportion of units in tract which are single family dwellings.
31. PGT5: Proportion of units in tract which are structures housing more than five units.
32. TRACT Y: Median income of tract

Appendix II. Coefficient Estimates for the Hedonic Rent Equations

	(1) Chicago		(2) Los Angeles		(3) Detroit	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
CONSTANT	3.7107	0.0637	3.5540	0.0922	3.4634	0.0647
1. 50–60	0.3354	0.0533	0.3336	0.0294	0.0737	0.0418
2. 40–49	0.2374	0.0456	0.2085	0.0357	0.1674	0.0265
3. 30–39	0.0742	0.0304	0.1135	0.0355		
4. PRE1920	−0.0831	0.0168			−0.0450	0.0152
5. 2RM	0.1586	0.0306	0.1087	0.0494	0.0590	0.0449
6. 3RM	0.3346	0.0322	0.2551	0.0506	0.1536	0.0438
7. 4RM	0.4472	0.0324	0.4017	0.0566	0.2416	0.0468
8. 5RM	0.5968	0.0342	0.4828	0.0656	0.3050	0.0464
9. MT5RM	0.6722	0.0351	0.7743	0.0867	0.3820	0.0483
10. BATHS	0.0840	0.0451			0.0783	0.0570
11. NOBATH	−0.0745	0.0259			−0.0343	0.0316
12. HHDUR	0.2500	0.0258	0.1352	0.0279	0.1371	0.0221
13. MA	0.0301	0.0100	0.0575	0.0240	0.0591	0.0108
14. NOWATER	−0.5320	0.0617				
15. NONCENT	−0.3096	0.0239	−0.0802	0.0391	−0.2539	0.0261
16. NONINST			−0.1502	0.0463		
17. GARAGE	0.0582	0.0361	0.0636	0.0297	0.0313	0.0165
18. GARAVAIL			0.0955	0.0349		
19. CAC	0.0255	0.0240	0.0563	0.0485	0.1467	0.0351
20. ELEV	0.1821	0.0314	0.0402	0.0296	0.0481	0.0253
21. PER/RM	0.0383	0.0178	0.0706	0.0222		
22. RACE	0.1856	0.0210	0.0912	0.0340	0.0913	0.0153
23. PLACK	−0.2462	0.0425			−0.1496	0.0462
24. PCRWD	−0.1867	0.0968	−0.1926	0.1887		
25. PGT5	0.2394	0.0321			0.1802	0.0266
26. PSFU			−0.3488	0.0477		
27. TRACT Y	0.0270	0.0060	0.0750	0.0010	0.0890	0.0070
R^2	0.6876		0.7957		0.6272	
SEE	0.2019		0.1495		0.1761	
DW	1.60		1.64		1.65	
F	85.0		47.1		70.7	
Mean Rent	91.3126		77.9300		74.7921	
Mean Log Rent	4.4520		4.3046		4.2728	
No. Obs.	952		276		861	

COEFFICIENT ESTIMATES FOR THE HEDONIC RENT EQUATIONS

	(4) Boston		(5) Pittsburgh		(6) Cleveland	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
CONSTANT	2.9726	0.1006	3.7397	0.1010	3.5408	0.1001
1. 50–60					0.3918	0.0744
2. 40–49				0.0693	0.2885	0.0520
3. POST1940	0.3989	0.0590	0.2042			
4. 30–39	0.0777	0.0337	0.1335	0.0755		
5. PRE1920					−0.0504	0.0187
6. 2RM	0.3339	0.0547	0.2168	0.0886	0.2699	0.0729
7. 3RM	0.4026	0.0528	0.3480	0.0880	0.3710	0.0707
8. 4RM	0.4923	0.0588	0.4821	0.0916	0.4416	0.0722
9. 5RM	0.5611	0.0669	0.5455	0.0960	0.5347	0.0732
10. MT5RM	0.5982	0.0673	0.6878	0.1064	0.6472	0.0755
11. BATHS			0.1481	0.0883		
12. NOBATH	−0.2292	0.0445	−0.1160	0.0334	−0.1192	0.0396
13. HHDUR	0.2700	0.0664	0.2824	0.0503	0.1320	0.0298
14. MA	0.0703	0.0332	0.0868	0.0370	0.0404	0.0160
15. COND			−0.0490	0.0366	−0.0466	0.0253
16. NOWATER	−0.2664	0.0629	−0.1182	0.0596		
17. NONCENT	−0.3685	0.0375	−0.1917	0.0335	−0.2068	0.0260
18. NONINST			−0.1513	0.0560		
19. GARAGE	0.0967	0.0602	0.0838	0.0776	0.0687	0.0217
20. GARAVAIL	0.0671	0.0592			0.0453	0.0271
21. CAC			0.1590	0.0711		
22. ELEV	0.0525	0.0373	0.3475	0.0744	0.1152	0.0504
23. PER/RM	0.0965	0.0352	0.0902	0.0373	0.0546	0.0210
24. RACE			0.1563	0.0377	0.1189	0.0216
25. PLACK	−0.2164	0.0507	−0.1089	0.0335	−0.0572	0.0545
26. PCRWD			−0.9542	0.2136		
27. PGT5	0.3754	0.0529	0.2092	0.1020	0.2258	0.0548
28. PSFU					0.1676	0.0685
29. TRACT Y	0.1190	0.0150			0.0310	0.0090
R^2	0.8258		0.7147		0.5366	
SEE	0.1828		0.2223		0.1897	
DW	1.56		1.64		1.50	
F	66.6		35.4		30.3	
Mean Rent	82.6485		69.6425		84.0581	
Mean Log Rent	4.3342		4.1639		4.3955	
No. Obs.	287		349		598	

Coefficient Estimates for the Hedonic Rent Equations

	(7) Washington		(8) Baltimore		(9) St. Louis	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
CONSTANT	3.6201	0.0871	3.1514	0.0796	3.3990	0.1149
1. 50–60	0.0550	0.0249	0.2788	0.0532	0.1596	0.0874
2. 40–49	0.0465	0.0214	0.0846	0.0327	0.1059	0.0503
3. 30–39			0.0682	0.0404		
4. PRE1920	−0.0264	0.0224	0.0670	0.0283	−0.0498	0.0230
5. 2RM	0.1817	0.0360	0.2007	0.0590	0.2614	0.0733
6. 3RM	0.3085	0.0343	0.3834	0.0556	0.4414	0.0728
7. 4RM	0.3999	0.0361	0.4823	0.0564	0.5292	0.0738
8. 5RM	0.4895	0.0407	0.6303	0.0597	0.6442	0.0765
9. MT5RM	0.6541	0.0655	0.7895	0.0684	0.7378	0.0795
10. BATHS	0.1669	0.0479	0.2099	0.0456	0.0853	0.0503
11. NOBATH	−0.0792	0.0387	−0.1388	0.0375	−0.1196	0.0264
12. HHDUR	0.1068	0.0283	0.2245	0.0271	0.2649	0.0317
13. HHDUR*			0.1139	0.0497		
14. MA			0.0594	0.0138		
15. COND					−0.0335	0.0232
16. NOWATER			−0.1445	0.0494	−0.1937	0.0279
17. NONCENT			−0.1268	0.0319	−0.2217	0.0244
18. NONINST					−0.2229	0.0433
19. GARAGE	0.0934	0.0421			0.0718	0.0260
20. GARAVAIL	−0.0621	0.0395				
21. CAC			0.0546	0.0292		
22. CAC*	0.2480	0.0379				
23. ELEV	0.2028	0.0247	0.2395	0.0394		
24. PER/RM	0.0735	0.0204	0.1223	0.0226	0.0370	0.0173
25. RACE	0.0294	0.0209	0.1607	0.0218		
26. PCRWD	−0.3004	0.1032			−0.1415	0.1173
27. PNWU					0.1262	0.0306
28. PSFU	0.2910	0.0988				
29. PLACK			−0.2358	0.0638	−0.1463	0.0458
30. PGT5	0.3531	0.0923	0.1645	0.0481	0.2507	0.0516
31. TRACT Y	0.0180	0.0040	0.0680	0.0070	0.0580	0.0130
R^2	0.5816		0.6830		0.7744	
SEE	0.1726		0.1947		0.2058	
DW	1.65		1.63		1.62	
F	40.7		59.3		96.9	
Mean Rent	89.1046		77.6572		63.5704	
Mean Log Rent	4.4529		4.2915		4.0626	
No. Obs.	637		657		644	

COEFFICIENT ESTIMATES FOR THE HEDONIC RENT EQUATIONS

	(10) San Francisco	
	Coefficient	Standard Error
CONSTANT	3.7490	0.1174
1. 50–60	0.3565	0.0809
2. 40–49	0.1361	0.0847
3. 30–39	−0.1107	0.0569
4. PRE1920	−0.1470	0.0282
5. 3RM	0.1739	0.0352
6. 4RM	0.2851	0.0410
7. 5RM	0.4115	0.0479
8. MT5RM	0.5250	0.0633
9. BATHS	0.1667	0.0630
10. NOBATH	−0.1725	0.0469
11. HHDUR	0.1541	0.0368
12. MA	0.0694	0.0182
13. NONCENT	−0.0616	0.0327
14. NONINST	−0.1469	0.0444
15. GARAGE	0.0571	0.0384
16. GARAVAIL	0.1082	0.0683
17. ELEV	0.0665	0.0384
18. PER/RM	0.0750	0.0336
19. PLACK	−0.3023	0.0797
20. PCRWD	−0.7200	0.2266
21. PSFU	−0.2140	0.0918
22. TRACT Y	0.0740	0.0160
R^2	0.6465	
SEE	0.2371	
DW	1.75	
F	31.6	
Mean Rent	84.5122	
Mean Log Rent	4.3575	
No. Obs.	403	

BIBLIOGRAPHY

[1] Gillingham, R., "Place to Place Rent Comparison Using Hedonic Quality Adjustment Techniques." Unpublished Ph.D. Dissertation, University of Pennsylvania, 1973.

[2] Gillingham, R., "Place to Place Rent Comparison Using Hedonic Quality Adjustment Techniques." Paper Presented at the NBER Conference on the Analysis of Consumer Expenditure Data, Palo Alto, May 2–3, 1974. Forthcoming as BLS Staff Paper.

[3] Glejser, H., "A New Test for Heteroscedasticity," *Journal of the American Statistical Association*, Vol. 64, (March 1969), pp. 316–23.

[4] King, A. T., "The Demand for Housing: A Lancastrian Approach," National Bureau of Economic Research, Conference on Research in Income and Wealth (Vol. 40), *Household Production and Consumption*.

[5] Lamale, H. H., "New BLS Standard Budgets and Living Cost Indexes," presented at the 45th Annual Agricultural Outlook Conference, Washington, D.C., November 1967.

[6] Lamale, H. H. and M. S. Stotz, "The Interim City Worker's Family Budget," *Monthly Labor Review*, Vol. 83, (August 1960), pp. 785–808.

[7] Lancaster, K. J., "A New Approach to Consumer Theory," *Journal of Political Economy*, Vol. 74, (April 1966), pp. 132–57.

[8] Lancaster, K. J. "Change and Innovation in the Technology of Consumption," *American Economic Review*, Vol. LVI, (May 1966), pp. 14–23.

[9] ———, *Consumer Demand*. New York: Columbia University Press, 1971.

[10] Muellbauer, J., "Household Production Theory, Quality and the Hedonic Technique," Birbeck Discussion Paper No. 1, University of London, 1973. Forthcoming in the *American Economic Review*.

[11] Pollak, R. A., "Subindexes of the Costs of Living Index," Research Discussion Paper No. 15, Research Division, Office of Prices and Living Conditions, Bureau of Labor Statistics, June 1973.

[12] ———, "The Theory of the Cost of Living Index," Research Discussion Paper No. 11, Research Division, Office of Prices and Living Conditions, Bureau of Labor Statistics, June 1971.

[13] Prais, S. J. and H. S. Houthakker, *The Analysis of Family Budgets*. London: Cambridge University Press, 1971.

[14] Rósen, S., "A Theory of Hedonic Prices," Department of Economics Discussion Paper, University of Rochester.

[15] Triplett, J. E., "Consumer Demand and Characteristics of Consumption Goods" BLS Working Paper No. 22, February 1974. Forthcoming in National Bureau of Economic Research, Conference on Research in Income and Wealth (Vol. 40), *Household Production and Consumption*.

[16] U.S. Bureau of the Census, *U.S. Census of Housing: 1960. Vol. I, States and Small Areas. United States Summary*. Final Report. HC(1)–1. U.S. Government Printing Office, Washington, D.C., 1963.

[17] ———, *U.S. Census of Housing: 1960. Vol. III, City Blocks*. Series HC(3), Nos. 46, 67, 105, 133, 178, 180, 204, 232, 309, 345. U.S. Government Printing Office, Washington, D.C., 1962.

[18] ———, *U.S. Censuses of Population and Housing: 1960. Census Tracts*. Final Report PHC(1), Nos. 13, 18, 26, 28, 40, 82, 118, 131, 137, 166.

[19] U.S. Bureau of Labor Statistics, *Three Standards of Living for an Urban Family of Four Persons*, Spring 1967. Bulletin No. 1570–5.

[20] Zar, J. H., "Significance Testing of the Spearman Rank Correlation Coefficients," *Journal of the American Statistical Association*, Vol. 67, (September 1972), pp. 578–80.

Annals of Economic and Social Measurement, 4/1, 1975

COMMENT

BY WILLIAM E. ALEXANDER*

Robert Gillingham is to be congratulated for the successful execution of his study, and I thank him for choosing as his subject the determination of multiple unit apartment rents, thereby relieving me of the task of discussing yet another hedonic study of the price determination of automobiles, or at least of some other durable good on wheels. In case anyone thought otherwise, Gillingham has successfully shown, in my opinion, that the technique of hedonics has a rather wider field of application than a survey of existing literature might suggest.

The physical magnitude of this study is appalling, especially if one is asked to review it subject to a reasonable time constraint. In excess of 5,600 observations have been moulded into 45 regression equations while these equations in turn have been distilled into 24 separate rent indexes. Furthermore, the text makes it clear that the reported results are only the visible portion of the iceberg. Fortunately for all concerned, since I cannot seriously quarrel with the empirical results presented here, I will endeavour to limit my remarks in this respect, and to concentrate upon what I regard the implications of this study to be for the future of hedonic research.

Essentially, Gillingham uses the hedonic equation as a form of specification pricing, presumably because it can handle the linking problems associated with non-overlapping market baskets. In his words, the object is to predict "the price of a unit which exists but is not observed" in the sample drawn. The unobserved unit in this case is the average apartment observed in some other comparison city. In order to carry this comparison out, Gillingham regards the attainment of hedonics equations which are as completely specified as possible as his major challenge, since he believes that misspecified proxy relationships are the bane of existing studies. He attacks in two ways: (1) by extending the list of quality determinants to include the location of the apartment; and (2) by careful attention to the functional form. I believe that he would regard these factors as those which most distinguish his study from the hedonic endeavours of others.

The attempt to specify the determinants of quality lead to two competing hypotheses: (a) that observed rent is related to the physical characteristics of the apartment, such as its age, its size (as measured by the number of rooms) and the presence or absence of a progression of luxuries such as furnished or otherwise, presence of hot water, bathrooms, garage facilities, etc.; (b) that in addition quality is determined by the physical condition of the neighborhood (as measured by proportion of block lacking plumbing facilities, proportion of observations in large buildings) and socio-economic status of the neighborhood (as measured by median income of the census tract). In addition, two variables signifying the race of occupant and the number of occupants per room are included. With the exception of these latter two variables, all determinants of "quality" may be given a

---

* The opinions expressed herein are the personal opinions of the author and no responsibility for them should be attributed to the Bank of Canada.

demand interpretation in the sense that their presence would be regarded as desirable by the consumer. The positive coefficient on the density variable (people per room) requires an unambiguous supply interpretation as a charge by the supplier for wear and tear associated with intensity of use of the structure. Gillingham is thus eclectic in his interpretation of the reduced form hedonic equation.

Generally, I am impressed with the results of these equations. The magnitude of the coefficients on the whole are acceptable on *a priori* grounds, although based on statistical significance, it appears that age and number of rooms do the bulk of the work. Undoubtedly sampling distribution has something to do with the robustness across cities of some of the other results. It is interesting, for example, that over 50 per cent of the apartments sampled were constructed prior to 1920 and that only 6.2 percent were less than ten years old, the latter varying from under 1 percent in St. Louis to over 25 percent in L.A. This apparently has led to substantial experimentation to find the optimum dummy classification for each city, (since final reported classifications are by no means constant across cities).

The results obtained by adding "location" variables are simultaneously encouraging and disappointing. They are disappointing in that the R^2 improvement is only about 6 points; however they are encouraging in that Gillingham's results suggest that exclusion of location variables does not appear to alter radically the coefficients of the included variables (constant term excepted). This is encouraging since we may infer that earlier studies excluding these variables are not deficient; it is also encouraging since it is an extremely difficult procedure to obtain this information (witness the necessity of combining separate surveys). I find the results of their inclusion plausible: the proportion of 5-unit buildings interact with the elevator variable, substandard housing reacts with race of occupant, and socio-economic location seems to interact with age. What is surprising to me is that only socio-economic location (as measured by tract median income) is consistently significant, and there is a possible interpretation problem here. That is, does tract median income measure location, or is it really measuring an income effect and asserting that rental housing is not an inferior good? Since other variables were available, such as tract median education, I would be interested in the unreported results for these variables.

Referring now to the specification of the functional form, there is little to say since there is no effective empirical distinction between the log-linear and linear form. The log-linear equation is favoured by Gillingham since it appears to exhibit the least heteroskedasticity. Personally, I am not surprised by the evidence of the presence of heteroskedasticity. However, as I shall presently argue, there may be superior methods of dealing with this problem.

Turning briefly to the actual rent indexes, we are met with a mass of apparent contradictions. Great differences appear in the ten city indexes, both in terms of dispersion and in terms of rank ordering. Yet I believe there is room for optimism: the hedonic technique does yield an approximate form of transitivity in that it consistently knows which are the cheap and which are the expensive cities. Detroit and Baltimore each get 9 out of 10 possible votes as one of the three cheapest cities, while San Francisco gets 6. L.A., Washington and Chicago get 7, 7 and 6 votes respectively as the three most expensive cities. And in many cases the differences among the middle-priced cities are extremely small. (This is not to deny the existence of anomalies, however. For example, Washington, consistently ranked as an expensive city, is ranked cheapest in terms of its own index!!!), a result I find puzzling. Yet in summary, I find the results encouraging, and I am not dismayed by the results of comparing hedonic indexes with the published indexes.

This statement begs the questions that, since simple indexes broadly confirm the hedonic results, *are* hedonic indexes worth the effort required to construct them or, can they be *made* worth it? This brings me to what I regard as the major shortcoming of this paper, and it is appropriate to raise it at this time because I feel that the problem lies partly in our failure to generate consumer expenditure data adequate to the needs of hedonic studies. Bob Gillingham began this paper by appeal to Lancaster's "New Theory" of demand as a means of justifying the hedonic approach to the measurement of quality. In any case, this historical *ex post* justification currently is being canonized in the literature. Gillingham rightly acknowledges that there are *difficult* problems of interpretation posed by this approach. But even so, in my opinion, the Lancaster theory does have explicit empirical implication for the conduct of hedonic studies, and I believe that they have had to be ignored in this study. For example:

(1) The choice of the "Characteristics" is arbitrary.

What are the relevant characteristics? The method of selection in this study (as in all hedonic studies) is a mixture of the author's priors ("I *think* it is a characteristic...") and *ex post* statistical verification ("it *must* be a characteristic if its coefficient has the expected sign and is larger in magnitude than its standard error..."). This procedure is totally inadequate. Lancaster has suggested that characteristics will be "revealed relevant." Yet I am not sure whether this approach will prove useful when we are forced to make observations at the *market* level for a complicated commodity like a multi-unit apartment dwelling. It is quite conceivable, for example, that rental housing can be a normal good for an individual at some points in his life cycle and an inferior good at other points. At these different points, his evaluation of the relevant characteristics might differ substantially. (Would you by choice raise children in a high rise?) Could these results be inferred from market data? Also, might an individual's perception of the characteristics of multi-unit apartments depend on his existing portfolio? The summer cottage, the ski chalet? They are ignored here. What is logically prior to the hedonic study is a carefully articulated survey of what people perceive the relevant characteristics to be.

(2) The possibility of distinct consumer groups existing in the same market and simultaneously reacting to different sets of implicit prices.

Lancaster's theory suggests that if different consumer groups exist, (different in the sense that their tastes are different or at least nonhomothetic), then we should not expect to find all the consumers in a market reacting to a single set of implicit prices unless the production technology of combining groups of characteristics exhibits constant returns to scale. In other cases, if linear combinations are allowed, for a given expenditure the set of consumable characteristics is a convex polytope and different consumers will be at equilibrium on its various facets. If such a model accurately depicts reality, then it is wrong to fit a single regression to all the observed data. What obtains is a weighted average of the facets, and its stability will depend on the stability across samples of the relative weights of the consumer groups on each of the facets. Furthermore, the problem will continue to exist even if hedonic regressions are looked upon as nothing more than reduced form equations. I believe that such a model is relevant in a segmented market like rental-housing. The implicit prices that the Park Avenue resident pays for a doorman and building security are not relevant to the housing tenant in a Detroit slum worried about rats and the presence or absence of running water, and it is wrong to include them in the same sample. Gillingham explicitly recognizes this taste problem at the end of his paper, and uses it as a possible explanation of the differences in city-by-city rent comparisons. However, I don't believe that geographical partitions of the sample adequately capture the taste problem described here. Again, what is needed is comprehensive socio-economic survey data generated at the individual observation level which would allow isolation of separate consumer groups and which was unavailable to Gillingham. (I note in passing that I have had some success along these lines using survey data on automobile purchases but once again, the occasions for obtaining such data are extremely rare).

(3) Choice of Functional Form

Contrary to Gillingham's position, Jack Muellbauer has pointed out that if one accepts the Lancaster model as the basis for hedonic studies, then the semi-log functional form will never obtain and instead, it is likely to be linear. However, it must be pointed out, that if the previous argument relating to distinct consumer groups is accepted, a strong case can be made for the semi-log forms. The linear form will be heteroskedastic (as Gillingham found). The superior approach would involve isolating separate consumer groups and fitting linear regressions corresponding to each facet of the characteristics possibilities set. Failing that, a nonlinear function may afford a reasonable approximation. This, I think, is how Gillingham's heteroskedasticity result is best interpreted.

In summary then, we would point to the need for a panel study carefully articulated to the needs of hedonic studies. In the meantime, however, Robert Gillingham's paper is in my opinion a fair expression of how well we are likely to be able to do until we get that data.

Bank of Canada,
Ottawa, Canada