Housing Allowance Demand Experiment

Hedonic Indices as a Measure of Housing Quality

Sally R. Merrill

June 1980

Abt Associates Inc., Cambridge, Massachusetts

ABT ASSOCIATES INC
55 WHEELER STREET, CAMBRIDGE, MASSACHUSETTS 02138
TELEPHONE • AREA 617-492-7100
TELEX 710-3201328

Submitted to:

Office of Policy Development and Research
U.S. Department of Housing and Urban Development
Washington, D.C.

HEDONIC INDICES AS A MEASURE OF HOUSING QUALITY

Contract H-2040R Task 3.3.3
Task 3.2.2

December 1977
(Revised June 1980) AAI# 76-96R

Principal Author: Sally R. Merrill

Contract Manager	Quality Control Reviewer	Management Reviewer

ABSTRACT

The major purpose of this paper is to develop an important summary measure of housing based on hedonic indices. The analysis of changes in housing is a central issue in the Demand Experiment. There is, however, no generally agreed upon measure of housing and housing quality. This paper assesses alternative measures of housing and proposes a general measure of housing based on estimated hedonic indices for use in later analysis.

Hedonic indices relate rent to measures of housing characteristics and thus provide one way of aggregating the many characteristics into an overall index of quality. There are several reasons why this approach is especially useful to the analysis of the Demand Experiment.

> First, an hedonic index can incorporate a wide range of attributes into one measure of housing, including not only the quality and size of the dwelling unit, but also many characteristics of the neighborhood, such as the quality of its housing stock and the quality of public services provided.
>
> Second, the hedonic index provides a stable and reasonable measure of housing over the two-year experimental period. The change in the housing index is adjusted for a variety of factors which may affect rent, including inflation, discounts for long tenure, or changes in landlord/tenant relations.
>
> Third, the hedonic approach permits investigation of a number of topics which bear on the analysis and understanding of the Demand Experiment. These include price discrimination against residents of minority neighborhoods, other types of housing market segmentation, and factors which result in some households getting a "better deal" for their money.

The derived hedonic indices presented in this report are based on evaluations of individual units by site office staff, participant ratings of their neighborhood, and other Census and local government data. They account for from 66 to 80 percent of the variation in rent and confirm the importance of dwelling unit and neighborhood amenities, as well as other nonquality characteristics, such as length of tenure, in determining market rent.

ACKNOWLEDGEMENTS

This research would not have been possible without the contribution of many people. A special debt is due Stephen Kennedy, Project Director, who co-authored the exploration of the theoretical foundations of the hedonic model in Appendix I, which helped to structure much of the discussion in the report, and especially Chapter 4.

Also, special thanks are owed to Cathy Joseph who was responsible for organizing and managing the massive and complicated data base required for this work. In addition to her work in developing many of the variables and computer requests, she carried out the analysis of adjustments to rent (Section 3.4 and Appendix III), and organized the appendices documenting the variables (Appendices VI and VII).

Walter Stellwagen, principal reviewer for the Demand Experiment, has painstakingly reviewed numerous drafts. He is responsible for many corrections and improvements in the conceptualization of the indexing problem and in the statistical approach and presentation of findings. James Wallace, Director of Design and Analysis, who had reviewed many earlier versions of this report, once again provided help for the final version.

The analysis benefited greatly from the comments made by the HUD review panel. In particular, Thomas King and Zvi Griliches provided many helpful suggestions concerning both test strategy and the interpretation of estimated hedonic indices.

Joseph Fisher contributed statistical expertise and in particular helped address the issues of collinearity and test level in Chapter 2. David Napior and Walter Stellwagen are responsible for the discussion of principal components analysis in Appendix VII.

Thanks also go to Irma Rivera-Veve who ably performed the very extensive computer work aided by Fred Luhmann and Nouna Kettaneh, Data Processing Manager for the Demand Experiment. The numerous tables were prepared by Diane Talamas, who also did the computations for the subsample housing bundles.

Finally, special thanks to Helen Bakeman, Deputy Contract Manager, for her efficient and persistent efforts to coordinate, schedule, and organize this research. And to Billie Renos, who typed numerous drafts and managed the final production of the report.

Sally R. Merrill

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF FIGURES

SUMMARY

Appropriate measures of housing are essential to address most of the policy concerns in the Demand Experiment, including the design of an effective housing allowance and the comparison of the housing allowance, other housing programs, and general income maintenance. Some of the major analytic issues being addressed include:

the extent to which the allowance is translated into improved housing;

the extent to which the relative cost of housing and the amount of housing purchased vary among different demographic groups or different experimental programs;

the extent to which the cost of housing obtained under housing allowance programs differs from that in existing housing programs.

This paper describes the development of an important summary measure of housing: a housing index derived from an hedonic model of housing. In addition, the model has been used to explore several topics pertinent to greater understanding of the experimental results such as price discrimination against minorities, housing market segmentation, and the discount associated with long-term residence in a rental unit. The actual application of the housing index to measure housing change in the experiment will be described in other reports.

A house or apartment is a complex bundle of attributes including not only the attributes of the unit itself, but also its neighborhood, and the quality of public and private services available. Any approach to measuring housing involves a decision as to which attributes of the housing bundle should enter into the definition of housing and how they should be weighted in determining the overall level and quality of housing services. Since a means of combining and weighting attributes is not given, it must be derived from some external criterion, such as market value, consumer satisfaction, or a normative concept of adequacy. The approach to these issues depends on whose point of view is being considered and what use is to be made of the quality measure. Policymakers, health and safety planners, environmentalists, individual

consumers, and the "marketplace" may each select and weight housing attributes quite differently. Thus, it should be emphasized that no single definition of quality is likely to be ideal for all purposes.

Measurement of Housing Quality

Three approaches to measuring housing quality will be used in the Demand Experiment: a measure of housing standards, reflecting policy concerns; individual preferences and satisfaction with their dwelling unit and neighborhood, reflecting the extent to which the unit conforms to the households' own needs and desires; and an index of housing services, reflecting market value. Each of these approaches addresses a unique aspect of housing quality. Used alone, none is ideal but each complements the others.

In terms of policy criteria, one of the experimental programs includes a Minimum Standards housing requirement, which participants must meet in order to receive an allowance payment. These Minimum Standards represent a modified subset of the American Public Health Association code and are very like the housing requirements used in the Section 8 Existing Housing Program. They thus provide a good proxy for a policy measure of acceptable housing.

At the same time, whether a unit passes or fails such standards provides a very limited measure of housing quality. Units are either acceptable or not; there are no gradations of quality. Furthermore, the measure is limited to a few features of the unit itself; it provides no indication of neighborhood quality, nor any reflection of recipient satisfaction with housing. Finally, the Minimum Standards themselves are not irrefutable. Any specific item may be challenged.

Another method for rating the quality of units is to use the tenant's expressed satisfaction with the unit and its neighborhood. Such measures are discussed in other reports (Weinberg et al., 1977; Atkinson and Phipps, 1977). These measures complement the Minimum Standards and the housing index by taking explicit account of recipients' own sense of the adequacy of their housing.

While all three measures are discussed, this paper is devoted largely to the hedonic index approach for measuring housing quality. The report applies the hedonic approach to data from Pittsburgh and Phoenix and explores the meaning and limitations of the resulting indices. In addition to separate housing indices for Pittsburgh and Phoenix, a common index which combines the data for the two cities has been developed to be used in certain analyses.

The hedonic approach assumes that, for the market as a whole, rent is strongly related to the quality and quantity of housing, in the sense that higher priced units reflect a general consensus that they offer more or better housing. However, rent is also determined by a variety of other factors not related to housing. Inflation, by definition, raises the dollar amount of rent without changing quality. Also, long-established tenants may pay lower rents because they are known to the landlord as good tenants, or because long residency may reduce landlord costs, or simply because it is easier for landlords to adjust rents upward when a unit is turning over. Racial discrimination may force minorities to pay more for comparable units. Individual households may simply obtain better deals, paying less than others for a given quality unit.

Hedonic indices essentially attempt to sort out the influence of housing and nonhousing factors in determining the market value of units (their rent). This allows the construction of indices which are sensitive to both unit and neighborhood characteristics and which do not include nonhousing factors such as inflation, tenure conditions, or racial discrimination.

Summary of the Results

The results of research done on hedonic indices, summarized below, describe both the characteristics of the hedonic model and the results of some of the applications of the model to assessing market segmentation, stability over the experimental period, and the extent of differences in the Pittsburgh and Phoenix housing markets.

1. The explanatory power of the hedonic estimates in Pittsburgh and Phoenix is quite high.

   In Pittsburgh, 66 percent of the variation in rent (and in the logarithm of rent) is explained by the available data; in Phoenix, nearly 80 percent of the variation is explained.

These results give some assurance that the hedonic approach provides reasonable measures of housing quality.

2. A large number of variables representing housing attributes are significant; furthermore, they represent attributes from all major component groups of the housing bundle.

   The significant variables represent dwelling unit facilities, dwelling unit quality, dwelling unit size, and neighborhood public services and amenities. Thus, the measure of quality will be sensitive to changes in the consumption of a very broad range of housing services. Moreover, if interest centers on the derivation of subindices representing dwelling unit quality or neighborhood quality, these subindices should also be sensitive to changes in many individual attributes.

3. Many tenure characteristics are significant in explaining variations in rent.

   The equations show that conditions of tenure (particularly length of residence, but also relationship to landlord, or presence of landlord in the same building) do, in fact, have an important effect on observed expenditures. This finding confirms the need to adjust for such factors when assessing changes in housing quality. For example, the estimated discount associated with a ten-year length of residence is about $15 per month in Pittsburgh and over $20 in Phoenix.

4. The hedonic model appears to provide stable and reasonable estimates over the two-year experimental period.

   The change in housing during the experiment is evaluated with reference to an index formed at the baseline period. This was done primarily because the baseline sample of households is by far the largest sample appropriate for estimation. In fact, the baseline model predicts at two years just as efficiently as a model estimated at the two-year period.

5. There is some evidence of price discrimination against residents of black submarkets in Pittsburgh.

   Extensive tests for price discrimination, on the basis of race of household and submarkets of different racial composition, were conducted in Phoenix and in Pittsburgh. Residents of ghetto areas in Pittsburgh--that is, in submarkets where more than 50 percent of the residents are black--appear to pay a price premium. It is small, however --about 4 percent. No evidence of price discrimination against either black households or Spanish American households was found in Phoenix.

6. **Housing market segmentation does not appear to pose any serious problems for the use of hedonic indices to measure housing quality.**

   To test for housing market segmentation, separate equations were estimated for central city and suburban areas as well as for racial submarkets. No evidence of central city/suburban segmentation exists in either Pittsburgh or Phoenix. While there is some evidence of segmentation between minority and white neighborhoods (Census tracts) and/or between minority and white households, it is minimal. The price markup found in Pittsburgh ghetto neighborhoods does not affect use of the index. Thus, it seems appropriate to use one housing index for all participants within each city.

7. **An index derived from Pittsburgh and Phoenix data combined provides an approximate way to make direct cross-site comparisons.**

   Housing outcomes within each site will generally be described using the specific index derived for the site. However, if a direct comparison is desirable, a common index is necessary. While the common index appears reasonable, the Pittsburgh and Phoenix housing markets are in fact different and a common index must be viewed as an approximate measure.

8. **Estimates of the rate of inflation in Pittsburgh and Phoenix confirm the need to adjust for inflation in evaluating the change in housing.**

   Two different estimates have been made in each city of the rate of inflation during the two-year experimental period. One is based on the hedonic estimates and one is based on the change in actual rent for households that did not move. Within each city these estimates are close to each other. The inflation rate in rental housing costs over two years (roughly 1973-1975) appears to be about 13 to 15 percent in Pittsburgh and 7 to 10 percent in Phoenix.

CHAPTER 1

INTRODUCTION

One of the major tasks of the Housing Allowance Demand Experiment is to compare the housing obtained by allowance recipients under different allowance programs with the housing obtained by nonrecipients. This, of course, requires the ability to measure housing to determine whether one house is better than another and by how much. But housing is a multidimensional concept. Many types of attributes define a house, and a given dwelling unit may be adequate in some respects, but not others. Overall comparison of different units requires that attributes somehow be aggregated into a summary measure. Unfortunately, there is no single, "objective" approach to defining and weighting housing attributes that satisfies all policy or analytical concerns. Different groups--such as individuals, policymakers, and the "market"--attach different importance to the various attributes. And, since these orderings are not generally observed, some external criterion--such as consumer satisfaction, social adequacy, or market rent--must be chosen in order to derive relevant weights. Thus, three major issues arise in defining an appropriate procedure for measuring housing:

> Which of the many dwelling unit and neighborhood attributes are relevant in describing housing?
>
> How should these attributes be defined?
>
> How can these attributes be combined into a summary measure --that is, since weights are not observed, how should they be derived?

Policymakers, for example, may focus on a selected number of dwelling unit characteristics that concern safety or adequacy. Attributes are frequently defined on a binary basis, that is, as "adequate" or "inadequate." Since different policymakers and planners have different concerns, however, many sets of attributes and many approaches to weighting these attributes may evolve.

Thus, for example, one of the programs tested in the Demand Experiment includes a specific set of Minimum Standards for housing, including physical standards for recipient dwelling units.[1] These standards are based in part on the American Public Health Association model housing code and resemble the standards adopted for the Section 8 Existing Housing Program. Yet although the Minimum Standards requirement or similar requirements can serve as a reasonable proxy for the policy adequacy of a unit, it would be difficult to argue that they are the single policy standard. Individual policymakers would undoubtedly quarrel with the inclusion or exclusion of certain requirements or with their relative importance.

Since the Minimum Standards requirement includes neither neighborhood characteristics nor certain dwelling unit attributes, measures based on them fall short of describing overall quality. In addition, because they are defined on a pass/fail basis, the standards are extremely insensitive; no improvement is recorded unless all items pass and major and minor repairs cannot be differentiated. Some of these disadvantages can be lessened by using the Minimum Standards components to form an index, such as the sum of the components that are passed by the dwelling unit. However, neighborhood attributes are still left out, and the equal weights given the various components in forming the index do not necessarily reflect any policy weighting of component importance.

Housing may also be measured in terms of individual household needs and satisfaction. In this case, the measure is idiosyncratic; each household defines its own needs and thus whether its housing is adequate. Without a general consensus concerning reasonable levels of housing, however, measures based on expressed household dissatisfaction may only reflect

---

[1] The definition of Minimum Standard housing is found in Appendix V. Standards for 15 attributes of the dwelling unit, such as plumbing, heating, electrical facilities, and interior and exterior quality are included. Minimum Standards are one part of the requirements used in the Housing Gap Minimum Standards program tested in the Demand Experiment. In order for households in this program to receive full payments, their dwelling units must meet all the physical requirements of Minimum Standards, as well as an occupancy requirement. The occupancy requirement is also described in Appendix V. Alternative standards levels are described in Abt Associates Inc., Working Paper on Early Findings, Cambridge, Mass., January 1975.

unreasonable desires or the inadequacy of a unit in terms of the household's unique circumstances and needs. Likewise, expressed satisfaction may be attacked as reflecting not only whether a unit is adequate, but also the extent to which the individual has simply given up his just hopes and aspirations in the face of apparently hopeless odds.

The credibility of individually based measures of housing can be improved by identifying some common consensus about what constitutes adequate or inadequate, or better or worse, housing. Thus, for example, expressed housing preferences and the housing actually purchased by different households may be examined to identify and weight some underlying set of basic housing attributes. However, given the wide variety of individual tastes and specific housing needs in terms of climate, household size, household composition, and lifestyle, it may be extremely difficult to identify such a consensus in a convincing way.

One approach to this problem is to regard a unit's market value (its rent) as reflecting, at least in part, a general consensus about the amount and quality of housing services that the unit provides. The problem in this case is to sort out the relation between rent and housing services in order to identify a general index of housing. Such measures are market specific; they do not claim to identify either underlying individual preferences or any long-run production function, and they may vary from city to city and over time. Within a specific market, however, they can provide the basis for a sensitive index of housing based on a wide variety of unit and neighborhood attributes.[1]

This report concentrates on the development of a broad index of housing based on market values. This is the hedonic approach to housing measurement, which assumes that the dwelling unit and its location (or neighborhood) may be defined in terms of various attributes and summarized by a particular weighted sum of these attributes. Since the weights cannot be observed directly, they are estimated by regressing rent on the various attributes. The estimated coefficients are then used to aggregate housing

---

[1]For a more complete discussion of the theory of hedonic indices, see Appendix I.

attributes into an overall index of housing services. In effect, the hedonic index of housing reflects the notion that, in the market as a whole, the relative rent commanded by a unit reflects the housing services that it provides. If rent were related only to housing services, then hedonic indices would be unnecessary; rent could be used to evaluate the unit directly. Relying solely on expenditures to measure quality poses certain problems, however. A discussion of these problems also indicates how an hedonic measure of quality will be useful in analysis.

Some portion of the variation in unit rents may be due to factors not directly related to housing services, such as inflation in housing prices, the possible rent premiums paid by recent movers, or other particular conditions of tenure, such as being related to the landlord. The result is that different amounts of housing may be purchased at a given level of rent. The hedonic regression attempts to sort out the influence of such nonhousing factors on rent.

Variation in rents may also reflect the existence of submarkets. The use of expenditures as a measure of quality assumes that the housing market is unified, so that all households have equal access to every unit. If the market is segmented into submarkets because of neighborhood characteristics, the socioeconomic preferences of residents, or racial discrimination, rent/quality differentials may exist, and individuals paying the same amount of rent may obtain different housing. The hedonic approach provides a method for assessing the extent of market segmentation. However, the existence of independent submarkets, in which the relative prices of attributes are markedly different, could undermine the use of a unified (market-wide) hedonic index by clouding interpretation of the housing attributes weights.

Finally, some households may be more able, or careful, or lucky, than others in shopping for housing. Thus individual households may achieve a given quality at a rent below the average market cost for that quality. In fact, differences in the incentives provided by the different programs tested in the Demand Experiment may add to existing differences in shopping ability or search behavior. For example, the Housing Gap Minimum Standards plan may encourage shopping practices different from the Minimum Rent or Percent of Rent plan to the extent that more care is taken to find housing meeting the

Minimum Standards at the lowest available rent.[1] The hedonic approach allows for an assessment of this behavior and more generally provides a means for adjusting individual differences by estimating a general market-wide relationship between rent and quality.

Several other analytical issues may be addressed by using hedonic measures of housing quality. For households that do not move, but whose rent increases, it is useful to determine whether any detectable change in quality accompanied the rent increase. This is particularly relevant in evaluating Minimum Rent households that do not move, but whose rent increases sufficiently to meet the Minimum Rent requirement. A housing index may also be useful in assessing the approximate value of changes for Minimum Standards households that upgrade and comparing this with any increases in rent.[2]

In addition, differences in the relationship between rent and quality may affect a household's level of satisfaction, its propensity to move, and possibly its desire to participate in the experiment. For example, households that get a better deal for their money may have less incentive to move or to participate in the program. Thus, hedonic measures of rent/quality ratios may be useful in analyzing the response of households to the housing allowance program.

Like all aggregate indices, hedonic indices may mask important information about specific aspects of housing change. In addition to examining changes in individual attributes, hedonic indices may also be used to form subindices. Thus, for example, if unit size, neighborhood characteristics, and other dwelling unit descriptors are all entered in the hedonic index, subindices may be formed for neighborhood and unit attributes or for unit quality in terms of a given unit size.

On the other hand, some quality features may have no rent value. For example, although high quality paint, copper water pipes and craftsmanship

---

[1] See Appendix IV for a brief description of the plans tested in the Demand Experiment.

[2] It is unclear whether hedonically derived indices will be sensitive enough to provide a complete analysis of such issues. Fortunately, other data, particularly survey responses, are available on the changes made in the units of households that do not move.

commonly command some market price in purchased housing, they probably do not affect rent. Returns to a landlord take place through lowered maintenance costs or longer building life, not directly through higher rent. Similarly, many features required by law, especially if the requirement varies by building size, may command no rent value and affect only profit and loss. Other items, valued by consumer and policymaker alike, such as the absence of lead-based paint, may have no effect on rent because they are simply not visible. If buyers are unaware of whether the hazard exists, they will not pay more to avoid it. Thus the hedonic index cannot fully replace more explicitly policy-oriented measures such as those based on the Minimum Standards used in the experiment.

Likewise, while the hedonic index may claim to reflect an overall market consensus about the amount and quality of housing services provided by a unit, it does not assess the extent to which a unit meets the needs of an individual household. Thus, in particular, a Minimum Standards requirement may require households to purchase units that meet Minimum Standards at the cost of sacrificing qualities of space or location that the household would prefer. Differences in expressed satisfaction, on the other hand, may provide a much better measure of the extent to which an individual household's housing preferences are met.

To summarize, an hedonic index attempts to provide an overall index of housing, independent of the effects of tenure characteristics, inflation, price discrimination, and shopping ability. It facilitates an analysis of housing change for households that do not move but upgrade their units, as well as for households that move. The relationship between the index and rent may help assess experimentally induced differences in shopping ability and may contribute to an analysis of satisfaction, moving behavior, and participation.

No general agreement exists, however, on a unique definition of quality, nor, from a policy perspective, on the "goodness" of dwelling unit attributes. An hedonic index, for example, cannot differentiate housing quantity and quality; rather, this distinction must be made by the household or a normative standard. Thus, in the analysis, the term housing quality, in the broad sense, includes several other measures, such as the

Minimum Standards, expressed satisfaction, measures of crowding, and individual attributes of the dwelling unit and neighborhood. The hedonic index is used in conjunction with these measures to provide a more complete picture of housing change.

The next four chapters (Chapters 2-5) consider the hedonic approach to quality measurement. This depth of coverage is necessary because the development and interpretation of an hedonic regression is complex. Chapter 2 describes the general hedonic approach to indexing housing, discusses the functional form of the equation, and outlines the regression strategy used to define a "good" equation. The estimated hedonic equations are presented in Chapter 3. Chapter 4 assesses the stability of the hedonic weights within and between the two Demand Experiment sites and over time. Finally, Chapter 5 illustrates the use of hedonic indices in considering the effects of racial and ethnic segregation.

CHAPTER 2

THE HEDONIC APPROACH

Empirically, the hedonic approach refers to systematic regressions of rent on housing characteristics. This approach has been applied to housing for a variety of analytical purposes.

Robert Gillingham (1975), for example, has used hedonic equations to construct price indices for housing across cities. Thomas King (1973) employed hedonic regressions to analyze the effects of property tax rates on housing price--that is, to determine the extent to which taxes are capitalized into location value.

Analysis of the effects of neighborhood amenities and public services on housing price and locational choice is frequently based on hedonic regressions. The measurement issues associated with defining appropriate neighborhood attributes have been analyzed by John Kain and John Quigley (1970) and by Thomas King (1973). Interest has centered both on general neighborhood quality and on specific attributes such as school and police quality. The effect of air pollution on property values has been estimated using hedonic regressions (Ridker and Henning, 1967).

Hedonic regressions have also been used extensively in analyses of price discrimination against minority households. To the extent that hedonic equations permit "standardization" of the housing bundle, the question of whether minority households pay more than nonminorities for equivalent housing can be addressed directly.[1]

Finally, much recent analysis has focused on determining the extent to which local housing markets are segmented into discrete submarkets, each with its own price structure. In order to assess the extent of market

---

[1]The effect of racial discrimination on housing prices has been analyzed in: Ridker and Henning, 1967; Haugens and Heins, 1969; Kain and Quigley, 1975; King and Mieszkowski, 1973; Quigley, 1974; Stengel, 1973; Schnare, 1974; Gillingham, 1975; Muth, 1969; Rapkin, 1966; Bailey, 1966; Lapham, 1971; Daniels, 1975; Berry and Bednary, 1975; Merrill, 1976.

segmentation, hedonic equations have been stratified by numerous neighborhood attributes, by tenure and structure type, by political jurisdictions, and by household characteristics.[1] In several cases, the regression coefficients have been used as attribute prices in estimating demand functions for housing attributes. Both King (1973) and Strazheim (1975), for example, have assessed the variation in demand for attributes in response to variation in their prices (see Strazheim, 1975; King, 1973).

The hedonic model estimated for housing is generally of the form

$$R = F(D,N,A,T,H). \tag{1}$$

Rent (R) is expressed as a function of dwelling unit characteristics (D), neighborhood attributes and amenities (N), access characteristics (A), and conditions of tenure (T). Frequently, household characteristics (H), such as race, are also included in order to test for price differentials based on these characteristics.

The determinants of rent in Equation (1) include items, such as tenure conditions or household characteristics, that do not involve the quality of the unit itself. Thus, if X_i is the vector of quality attributes for a particular unit, and Z_i is the vector of nonquality attributes, one can estimate

$$R_i = F(X_i,Z_i,e_i) \tag{2}$$

where e_i is stochastic and use $\hat{R}_i = \hat{F}(X_i,\overline{Z}_i,0)$ as an index of housing quality, where $\overline{Z}_i$ represents some fixed values for nonquality items.

Numerous questions must be addressed, however, in order to make the model operational. First, there are many potential measures of housing attributes, many of which are highly correlated with one another. As usual, some basic strategy must be adopted for variable definition, variable inclusion, and assessment of the remaining collinearity among variables. These issues are discussed in Section 2.1.

---

[1] See for example, Strazheim, 1973; Quigley, 1973; Schnare and Struyk, 1974; Kain and Quigley, 1975.

Second, theory provides little or no guidance as to the form of the hedonic function, except that it may often be nonlinear.[1] Section 2.2 discusses and compares the two functional forms dealt with in this report--a linear and semilogarithmic form.

Finally, the exact use of the hedonic index depends to some extent on the form adopted and on the interpretation developed for the estimated coefficients and residuals. This issue is discussed in Section 2.3.

2.1 SPECIFICATION OF THE HEDONIC EQUATION

Derivation of the "best" or final form of the hedonic equation is a complex and often ad hoc empirical process. Neither the other types of models used in housing market analysis nor the general hedonic model provides much guidance in the selection or definition of appropriate variables. There are many potential variables and they are often highly correlated, so that empirical tests often do not readily distinguish among alternative subsets. This section describes the sample and basic procedure used in estimating the final hedonic equation.

Sample and Data Sources Used for Estimation

The sample used to estimate the hedonic equations is based on all households enrolled in the experiment. The sampling procedure used to select participants is described in detail elsewhere (Abt Associates Inc., 1973). Two summary comments are relevant, however. First, the sample does not represent a random selection of dwelling units; rather, it is a random and unstratified sample of renter households that meet certain eligibility requirements, primarily an income limit that varies with household size. Second, based on an examination of participant and Census data, the distribution of demographic characteristics in the sample is quite similar to the distribution of these characteristics in the city as a whole for households within the same income limits.[2]

---

[1] See Appendix I.

[2] For a complete description of the demographic characteristics of the participants, refer to Abt Associates Inc., Working Paper on Early Findings, Cambridge, Mass., January 1975.

The necessary data were collected using several different instruments. The dependent variable used in the hedonic regression is contract rent adjusted to exclude utility costs.[1] The housing attributes used in the regression were of several types: most physical descriptors of the unit were taken from housing evaluations performed by site office staff in Pittsburgh and Phoenix; data on participant perceptions of their neighborhoods and on some attributes of the dwelling unit were collected in the Baseline Interview; various neighborhood descriptors were taken from the 1970 Census.

The Baseline Interview generally preceded the Initial Housing Evaluation by three or four months.[2] To assure that rent and dwelling unit data pertained to the same unit, 227 households that moved between the Baseline Interview and Initial Housing Evaluation were excluded. In addition, a smaller number of households that had not moved were excluded because they lacked complete data. The net sample used for estimation comprised 1,615 households in Pittsburgh and 1,614 households in Phoenix.

Equations for Pittsburgh and Phoenix were estimated separately, retaining different variables in the two cities as appropriate. Differences between the two cities are reflected in the final equations reported in Chapter 3. Pooled estimates were also developed using a common variable list for both sites, as discussed in Chapter 4.

Regression Strategy

The four major criteria used in deriving the final equation may be summarized as follows:

- maximizing explanatory power, that is, maximizing adjusted $\bar{R}^2$;
- including a broad set of attributes whose coefficients are significant at the chosen level and have the "expected" sign;
- defining variables such that the stability of the estimated coefficients does not appear to be unduly affected by collinearity; and

---

[1] This adjustment is further discussed in Chapter 3 and Appendix III.

[2] See Appendix IV for a summary of the data sources.

minimizing reliance on proxy relationships so that all major component groups, including dwelling unit quality, dwelling unit size, neighborhood quality, accessibility, and tenure relationships, are adequately represented in the equation.

Test Strategy

Preliminary equations were estimated using a large number of the derived variables in a given equation. The complete set of variables tested is listed in Appendix VI. Many of the coefficients were smaller than their standard errors or did not have the expected sign. Judgments about what to do with these variables were based on examination of the variance inflation factors, the correlation matrix, sample variances, and on intuitive decisions concerning the importance of the attribute.[1] Many variables whose coefficients were approximately equal to their standard error were kept for further consideration. Variables with the wrong signs, especially when the coefficient differed significantly from zero, were examined for evidence of incorrect definition or specification, or of extreme correlation with other variables in determining whether they should be redefined, combined with other variables, or simply dropped from the equation.

The basic goal of subsequent estimations was to weed out superfluous variables and to find the most broadly descriptive combination of housing bundle attributes. (A variable is defined as superfluous when its exclusion or inclusion in an equation does not "seriously" alter the other coefficient estimates or their standard errors (Rao and Miller, 1971).) Alternative definitions of variables were tested by using each definition in turn in the same regression equation and then adopting the definition that gave the smallest residual sum of squares (Rao and Miller, 1971).

A stepwise regression routine was used to a limited extent in deriving the best variable set.[2] Sole reliance on the stepwise regression results is inappropriate, since the procedure does not guarantee an optimal subset

---

[1] Variance inflation factors, used to assess collinearity, are defined below.

[2] The available program is basically a forward inclusion technique (Statistical Package for the Social Sciences, 1975). It first selects the variable most highly correlated with the dependent variable. The variable that explains the greatest amount of variance in conjunction with the first variable is entered next. Any level of significance for coefficients can be defined to limit the inclusion process.

(see, for example, Hocking, 1976). This routine was used primarily to search for the most significant neighborhood variables once an appropriate set of dwelling unit attributes had been derived.

The derivation of an appropriate set of neighborhood variables was generally much more difficult than the selection of a set of dwelling unit variables. In addition, correlations among the variables between these two attribute groups are generally lower than correlations within each attribute group --that is, the simple correlation between any given dwelling unit descriptor and any given neighborhood descriptor was generally quite low. Neighborhood and dwelling unit variables were therefore viewed as "distinct" sets. Tests to find the "best" set of dwelling unit variables were made using a tentative initial set of neighborhood variables; once dwelling unit descriptors were fixed, alternative groups of neighborhood variables were then tested.

The test used for the final hedonic regressions was a t-statistic of at least one. This implies a nominal level of significance of at least $\alpha = 0.25$ for a two-tailed test, and $\alpha = 0.125$ for a one-tailed test. (Since prior expectation as to sign exists for nearly all the coefficients, the one-tailed test is generally appropriate.)

The use of a less stringent test level reflects an emphasis on predictive power. The hedonic regression is used in analysis primarily to derive an overall estimated housing index; predictive power is thus of major importance. The adjusted $\bar{R}^2$ for an equation is improved by retaining any variable that has a t-statistic of greater than one (Haitovsky, 1969).[1] More generally, various authors have pointed out that the mean square error of prediction will be reduced if what might be called the theoretical F-statistic (the F-statistic using the true parameter values) is greater than one (see Edwards, 1969; Rao, 1971; and Hocking, 1974), which leads them to suggest a test level of one. (This applies to the t-statistic as well, since it is simply the square root of the F-statistic for one variable.) In addition, as Rao (1971)

[1] Adjusted $\bar{R}^2 = 1 - (1-R^2)(N-1)/N-K$, where N equals the total observations and K equals the total number of parameters. Adjusted $\bar{R}^2$ can decrease when a new variable is added to the regression, even though R^2 increases. Thus adjusted R^2 provides a check against the absurd extreme at which a "perfect" fit of N observations is obtained by the use of K = N parameters.

points out, omission of a relevant predictor will bias the included variables correlated with it, whereas inclusion of an irrelevant variable introduces no bias.[1]

Collinearity

Several groups of variables were very highly collinear. This was expected; better housing units are frequently better in terms of many attributes and are more often found in better neighborhoods. Any data set can be transformed into an orthogonal set without changing predictive power, so that multicollinearity does not affect the overall fit of the model.[2] It does, however, affect the reliability and stability of the coefficients. The more collinear the variables are, the larger is the variance of estimates of their coefficients. It therefore becomes more difficult to reject the null hypothesis that a coefficient is zero, and the coefficient estimates become less reliable. If collinearity is severe, some coefficients may, for example, reverse sign as a result of negligible changes in the data (Marquart and Snee, 1975).[3] Thus, although collinearity in the predictor matrix cannot be changed--this is a sample problem--it may be appropriate to respecify highly collinear variables to help derive a more interpretable set of parameter estimates.

This approach was followed in specifying the hedonic regressions. As discussed in Section 3.2, collinearity was particularly troublesome among variables describing the surface and structural quality of the dwelling unit, the characteristics of Census tracts, and residents' perceptions of their neighborhoods. Data reduction techniques for highly correlated variables were used in preference to excluding some subset of the original variables. Although the mean predictive power is not reduced by exclusion of

---

[1] These are heuristic arguments. It may be noted, for example, that the estimated F-statistic is not an unbiased estimator of the "theoretical" F-statistic.

[2] It may, however, affect prediction in new samples in which the set of associations do not resemble those in the sample used for estimation.

[3] Prediction in the directions represented by changes in one of the collinear variables, holding other variables constant, for example, has a correspondingly larger variance.

sufficiently insignificant variables, estimates of housing quality change for an individual household will be less accurate if the "excluded" variables were the ones that changed.

Data reduction techniques such as principal components analysis allow estimation of stable coefficients for the components, while retaining a large number of data elements within the component itself (see, for example, Cheng and Iglarsh, 1976).[1] These seemed more desirable than proxy variables, such as median Census tract income, which has frequently been used to represent neighborhood quality. It is true that this quantity is highly correlated with many attributes that define neighborhood quality, but when used as a proxy it provides no information about the relative importance of attributes. Thus a quality index that relies on income to measure neighborhood quality is less likely to be sensitive to real changes in quality than is an index based on specific neighborhood attributes.

The extent of collinearity in a regression model can be partially determined by assessing the simple correlation between any two predictor variables to see whether any of the relationships seem excessive. Frequently, however, the pattern of associations is more complex. Since linear dependency may involve several variables, a more comprehensive approach, based on the variance-covariance matrix of the estimated coefficients, is desirable.

The variance-covariance matrix of the estimated coefficients, $\hat{\beta}$, is given by:

$$U(\hat{\beta}) = \sigma^2 (X'X)^{-1} , \tag{3}$$

where

$\hat{\beta}$ = the vector of estimated coefficients

$(X'X)^{-1}$ = the inverse of the variance-covariance matrix of the independent variables, X

σ^2 = the variance of the error term in the regression, ε, where ε is distributed independently of X.

---

[1] In principal components analysis, the original variables are transformed into an equal number of new variables called principal components. These new variables are constructed so that they are not correlated. If the first few components have relatively large variances, these components are then used to summarize the original data. Appendix VII includes a discussion of this approach and a list of the variables so derived.

The variance of any particular coefficient, $\hat{\beta}_j$, is given by

(4) $$V(\hat{\beta}_j) = C_{jj}\sigma^2,$$

where C_{jj} is the j^{th} diagonal element of $(X'X)^{-1}$. If the variables, X, are standardized, C_{jj} is known as the variance inflation factor and becomes increasingly large as multicollinearity increases. For standardized variables, C_{jj} may be written as:

(5) $$C_{jj} = 1/(1-R_j^2),$$

where R_j^2 is the square of the multiple correlation between X_j and all the other predictor variables in X. If an independent variable is orthogonal to the remaining variables, R_j^2 is zero and the variance inflation factor is one. As R_j^2 approaches one, C_{jj} becomes increasingly larger. Variance inflation factors greater than four have been suggested as a rule-of-thumb indication that multicollinearity is excessive (Snee, 1970). When C_{jj} equals four, $R_j^2 = 0.75$; only one-fourth of the variability of X_j is orthogonal to the other predictors.

Appendix VIII lists the variance inflation factors, C_{jj}, for the final full sample equations discussed in Chapter 3. None is larger than 3.3 and most are less than 2.0. This reflects the use of various data reduction techniques, such as principal components, as well as the fact that more collinear variables have higher variances of estimate and are likely to be dropped from the estimating equation. In any case, the final equation does not seem subject to instability due to collinearity between included variables.

Omitted Variables

Omission of variables from the estimating equations increases the variance of estimate, and, if the omitted variables change from sample to sample, may cloud the interpretation of the hedonic residuals (in particular, the extent to which they reflect omitted housing attributes in addition to simple variation in the amount of rent paid for comparable units). Given the complexity of the housing bundle and the possible sensitivity of value to the exact arrangement of rooms, facilities, and so forth, of course, there will be omissions from the variables list. To some extent, inclusion of a wide range of both neighborhood and dwelling unit variables hopefully reduces the effect of such omission. In addition, however, further efforts were made to identify and, where possible, to reduce the effects of omitted variables.

The available neighborhood descriptors were generally felt to be less precise than dwelling unit descriptors. Given the size of the sample, there were severe limits to the level of resolution for neighborhood descriptors. Although ratings of the block face were available from the Initial Housing Evaluations, most descriptors were only defined at the level of the Census tract. This was particularly bothersome in Pittsburgh, which has a large number of well-defined neighborhoods. Plots of residuals on maps of the two sites were examined to identify any geographic groupings that might indicate the systematic effects of omitted neighborhood factors.

In addition, efforts were made to reduce the sensitivity of the estimated coefficients to extreme values of omitted variables. Extreme outliers were identified and the hedonic equations were reestimated without them in order to assess their influence on the estimates.

Finally, the estimated residuals were themselves regressed on various demographic variables. As discussed in Chapter 3, the results of these regressions at least suggest the nature of the variables included in the omitted items. Ill-conditioned residuals need not be solely due to omitted variables, of course. The form of the hedonic equation may also be misspecified. This is discussed further in the next section.

2.2 FUNCTIONAL FORM OF THE HEDONIC EQUATIONS

The general hedonic model does not dictate any particular functional form for the relationship between the market price of a commodity and its attributes. Most hedonic regressions for housing have used either a linear form or semilog form. The general linear form is

$$R_j = \sum_i \beta_i X_{ij} + e_j \tag{6}$$

for households, "j," and attributes, "i," and the semilog form is

$$\ln R_j = \sum_i \beta_i X_{ij} + e_j . \tag{7}$$

Four criteria were used to assess the linear and semilog functional forms: interpretation of the implied relationship, explanatory power, heteroskedasticity of the error term, and use of the housing services index in the analysis. Since, as detailed below, each form has particular advantages and disadvantages, both equations were retained for use in the analysis.

The first criterion is the a priori appeal of the implied relationship between rent and the attributes. The semilog model represents a multiplicative relationship--that is, it assumes that variables are jointly related to rent. Thus,

$$R_j = \exp\left(\sum_{i=0}^{N} X_{ij}\beta_i + e_j\right). \tag{8}$$

The coefficients of the variables are interpreted as the percentage of change in rent that results from a unit change in the level of the independent variable. The semilog form is appealing because it allows for a limited type of interaction among the variables. The implicit assumption is that the value of each attribute is a function of the overall amount of quality of the unit. The dollar value of having a garage, for example, is assumed to differ for a high quality and a low quality unit. In effect, the semilog form uses the overall quality of the unit to scale the quality of the garage. Since many variables are entered as (0,1) dummies without further rating of quality, this is a desirable feature. Most obviously, the semilog form allows the estimation of a multiplicative relationship between unit size and quality in which the absolute change in rent that results from an increase in average interior quality will be greater for larger units.

The linear form, on the other hand, allows for the explicit introduction of appropriate interdependencies. Some variables may be independently and others jointly related to rent (King, 1976). For example, it is unclear whether the value of neighborhood attributes is necessarily proportional to either dwelling unit quality or size (Grether and Mieszkowski, 1974). Some effort has been made to define variables that explicitly permit testing for nonlinear effects. For example, interaction variables were defined relating number of rooms and interior quality, rooms and building type, and interior quality and neighborhood quality. The results were negative. In some cases these interactions were clearly insignificant; in other cases the collinearity between the main effects variables and the interaction terms caused extreme changes in the coefficients, sometimes resulting in a reversal of sign. Use of the interaction terms alone did not improve explanatory power. Thus, either because the hypothetical interdependencies were not important or because there were statistical obstacles to estimation, the interaction terms have been dropped from the linear equations.

The second criterion for choosing between the linear and semilog forms was explanatory power. The percentages of variance explained, $\bar{R}^2$, of the linear form and semilog forms are nearly identical. However, this need not mean that the forms are equally powerful. Since the dependent variables are different, the residual sums of squares cannot be directly compared. One approach, developed by Box and Cox (1964),[1] uses the likelihood of the estimated equation to compare explanatory power. As used here, the test essentially asks whether transformation of the dependent variable (into log form) is empirically appropriate; if it is, the log likelihood function for the semilog equation will fall relative to that for the linear equation.

In order to make this comparison, the variables and number of estimated parameters in each equation must be the same. As described in Chapter 3, the final hedonic semilog and linear equations have somewhat different variable lists. Thus both a linear and semilog equation were estimated and a separate test of explanatory power was made for each variable list. For both variable lists in both cities the semilog form of the equation has somewhat better explanatory power than the linear form (see Table 2-1).[2]

---

[1] This reference was pointed out to us by Zvi Griliches.

[2] Statistical tests of the difference in the log likelihood functions for the linear and logarithmic form must be regarded as being at best suggestive. Following Box and Cox (1964), the general functional form being tested is

$$\frac{R^{\lambda}-1}{\lambda} + \lambda = X\beta + \varepsilon$$

where ε is distributed $N(0,\sigma^2)$. Note that this reduces to

$$R = X\beta + \varepsilon \quad \text{for } \lambda = 1,$$
$$\ln R = X\beta + \varepsilon \quad \text{for } \lambda = 0.$$

(footnote continued)

Table 2-1

COMPARISON OF LINEAR AND LOGLINEAR EQUATIONS

DEPENDENT VARIABLE	INDEPENDENT VARIABLE LIST TAKEN FROM: LINEAR EQUATION	INDEPENDENT VARIABLE LIST TAKEN FROM: SEMILOG EQUATION
	PITTSBURGH	
Rent	R^2 = .656 $\bar{R}^2$ = .648 L = −4725.43	R^2 = .656 $\bar{R}^2$ = .648 L = −4676.15
Natural Log Rent	R^2 = .660 $\bar{R}^2$ = .653 L = −4641.55	R^2 = .662 $\bar{R}^2$ = .654 L = −4631.46
Difference between log likelihoods	Δ = 83.88	Δ = 44.69
	$\chi^2_{.99(1)} = 6.63$	
	PHOENIX	
Rent	R^2 = .786 $\bar{R}^2$ = .783 L = −4866.62	R^2 = .786 $\bar{R}^2$ = .783 L = −4871.38
Natural Log Rent	R^2 = .796 $\bar{R}^2$ = .793 L = −4852.21	R^2 = .804 $\bar{R}^2$ = .800 L = −4816.82
Difference between log likelihoods	Δ = 14.41	Δ = 54.56
	$\chi^2_{.99(1)} = 6.63$	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

NOTE: Test due to Box and Cox, 1964.

Gillingham (1973) offers another criterion by which to judge the linear and semilog forms. A common type of heteroskedasticity may occur in which the standard error is proportional to the conditional expectation of the dependent variable. The semilog form corrects for this type of heteroskedasticity. Gillingham uses a modification of a test by Glejser (1969) in which the absolute value of the residual is regressed on the predicted value of the dependent variable, in this case rent or the natural log of rent. In addition, the squared residuals are regressed on the predictions of the dependent variables. Thus, the proposed regressions are:

$$
\begin{aligned}
(9) \qquad |\hat{e}|_{\text{linear}} &= \alpha + \hat{\beta}\hat{R} + \varepsilon \\
|\hat{e}|_{\text{semilog}} &= \alpha + \hat{\beta}\widehat{\ln R} + \varepsilon \\
e^2_{(\text{linear})} &= \alpha + \hat{\beta}\hat{R} + \varepsilon \\
e^2_{(\text{semilog})} &= \alpha + \hat{\beta}\widehat{\ln R} + \varepsilon .
\end{aligned}
$$

The predicted value is certainly an imperfect estimate of the conditional expectation. Nevertheless, a significant relationship is taken to imply that either the standard error or the variance of the residuals is proportional to the $X\hat{\beta}$ matrix. The results, shown in Table 2-2, do not indicate

(footnote continued)

Given any value of λ, the maximum log likelihood is given, as usual, by the OLS estimates of β, using the transformed variable, $(R^{\lambda}-1)/\lambda$. Thus

$$L(\lambda) = -\frac{1}{2} N \ln [(N-p)\hat{\sigma}^2(\lambda)/N] - (1-\lambda)\ln R,$$

where N is the number of observations, and $\hat{\sigma}^2$ the mean squared error under OLS estimation against the transformed variable.

Box and Cox suggest that an estimate of λ, $\hat{\lambda}$, may be obtained by finding the value of λ that maximizes $L(\lambda)$. An approximate $100(1-\alpha)$ percent confidence interval for $\hat{\lambda}$ is given by

$$L(\hat{\lambda}) - L(\lambda) < \frac{1}{2}\chi^2(\alpha)$$

(In the hedonic regressions, for one degree of freedom, χ^2 (0.01) is 6.63.)

Notice, however, that the comparison of $L(\lambda=1)$ and $L(\lambda=0)$ never deals with $L(\hat{\lambda})$. Thus, since $L(\hat{\lambda}) \geq L(\lambda=1,0)$, the evidence is sufficient to reject the hypothesis of a linear form ($\lambda=1$). This is not, however, a test of the semilog form ($\lambda=0$), nor can it be taken to indicate in any way that the true value of λ is closer to one than to zero (i.e., that the true form is closer to a semilog than a linear form).

Table 2-2

TESTS FOR HETEROSKEDASTICITY

DEPENDENT VARIABLE	EQUATION					
	LINEAR			SEMILOG		
	$\bar{R}^2$	Coefficient	t	$\bar{R}^2$	Coefficient	t
	PITTSBURGH					
$\lvert e \rvert$	.035	.086	7.67	.008	-.042	3.77
e^2	.030	.039	7.13	.011	-.021	4.33
	PHOENIX					
$\lvert e \rvert$	.006	.027	3.26	.05	-.075	9.44
e^2	.010	.018	4.10	.045	-.030	8.70

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

that the semilog form exhibits less heteroskedasticity than the linear form. The R^2 statistics for the semilog residuals are lower in Pittsburgh, but higher in Phoenix, than those for the linear residuals.

Perhaps the strongest rationale for retaining both the linear and the semilog form is related to the use of the hedonic regressions in the analysis. For much of the analysis, interest centers on the dollar value of the index, not on the logarithmic values.[1] Also, the residuals from the regressions will be used in various analyses. In some cases it is convenient to use the residuals valued in dollars, especially to the extent that the residual represents omitted housing variables and an estimate of their value is needed. On the other hand, the housing index derived from the semilog equation is used in the log linear demand function in an analogous manner to log rent. Despite the log likelihood test, since there is no clear choice between the two forms in terms of prior plausibility, explained variance, or error structure, both have been retained.

2.3 THE DERIVATION OF HEDONIC QUALITY MEASURES AND THEIR USE IN THE ANALYSIS

Hedonic indices are estimated in the Demand Experiment in order to form a broadly based index of housing that can be used to estimate the effects of the various experimental housing allowance programs. Appendix I discusses in detail the theory of hedonic quality indices. This section explicates their application.

Loosely speaking, the general rationale for hedonic indices as a measure of housing rests on two assumptions:

> First, in the market as a whole, cost does reflect the amount and quality of housing services offered by the unit. More expensive units generally reflect a common consensus that they are better.

[1] It may be preferable to derive the linear index from the linear equation, since exponentiating the logarithmic values gives an estimate of median rather than mean rent. The extent of bias is a function of the error of estimate of the observation. For a complete discussion of the problem of log linear bias see Appendix V of Friedman and Kennedy (1977), and the bibliography cited there.

Second, although true for the market as a whole, the above assumption is often not true for individual cases. Individual rents may reflect factors such as shopping ability, luck, conditions of tenure, and racial or other discrimination, as well as the unit's housing services.

Estimated indices attempt to distinguish the market valuation of housing services offered by a unit from the nonhousing factors that also influence rent.

The baseline hedonic weights are estimated from

(10) $$R_j = \hat{\beta}_0 + \Sigma X_{ij}\hat{\beta}_i + \Sigma Z_{ij}\hat{\gamma}_i + e_j,$$

where

R_j = rent or log of rent of the j^{th} unit, depending on the form used

β_0 = constant

X_{ij} = the amount of the i^{th} housing attribute in the j^{th} unit

Z_{ij} = the amount of the i^{th} nonhousing attribute in the j^{th} unit

e_j = error term.

Thus observed rent is estimated as a function of housing attributes, X, of other factors, Z, and of individual idiosyncracies, "e."

The housing index, $\hat{Q}_j$, is then the estimated market value of the housing attributes.

(11) $$\hat{Q}_j = \hat{\beta}_0 + \Sigma X_{ij}\hat{\beta}_i \qquad \text{(linear form)}$$

$$\hat{Q}_j = \exp[\hat{\beta}_0 + \Sigma X_{ij}\hat{\beta}_i]. \qquad \text{(semilog form)}$$

Similarly, the change in quality is indexed by

(12) $$\Delta\hat{Q}_j = \Sigma \Delta X_{ij}\hat{\beta}_i \qquad \text{(linear form)}$$

$$\frac{\Delta\hat{Q}_j}{\hat{Q}_j} = e^{\Sigma \Delta X_{ij}\hat{\beta}_i} - 1. \qquad \text{(semilog form)}$$

The housing index nets out the costs associated with nonhousing items. It reflects overall market value and thus corrects for individual luck or shopping ability. In addition, since the weights, $\hat{\beta}_i$, are estimated for a single time period, the index reflects that time period. Comparisons of the estimated hedonic value of housing over time will not be affected by inflation.

The hedonic equation may be used to form separate subindices of dwelling unit quality and neighborhood quality. For example, assume that a subset of characteristics, $X_1, \ldots, X_q$, pertains to a dwelling unit. A subindex of the change in dwelling unit features, $\hat{Q}_j'$, could be estimated by

$$(13) \qquad \Delta\hat{Q}_j' = \sum_{i=1}^{q} \Delta X_{ij}\hat{\beta}_i \qquad \text{(linear form)}$$

$$\frac{\Delta\hat{Q}_j'}{\hat{Q}_j} = e^{\sum_{i=1}^{q} \Delta X_{ij}\hat{\beta}_i} - 1. \qquad \text{(semilog form)}$$

Estimation of the absolute level of dwelling unit quality for the subindex requires allocation of the constant term among nonquality items, dwelling unit quality, and neighborhood quality attributes.[1] However, dollar changes in the linear form and proportional changes in the logarithmic form can be indexed independent of the value of the constant or other variables.

Although the housing change index defined in Equation (11) is plausible, it is subject to a variety of reservations. First, to the extent that the error term, e_j, in Equation (10) reflects the effects of omitted quality attributes rather than shopping ability or luck, the $\hat{Q}$ index will not capture these quality changes. This problem is reduced to the extent that the included variables, covering a wide range of attributes, successfully capture the effects of the omitted variables. As an alternative approach to a quality change index, one could theoretically correct for nonquality components of rent by subtracting them from actual rent, that is, by forming an adjusted rent index, $\hat{A}_j$:

[1] Interpretation of the constant term in hedonic equations is sometimes ambiguous. As the equation is specified here, the constant represents at least the omitted category of all dummy variables. It also represents some kind of "basic" dwelling unit, that is, a structure with walls, a roof, land, etc. For technical reasons, hedonic functions for housing probably cannot be defined at very low levels for all attributes. One possible solution is to allocate the constant term of a subindex in proportion to the share of that subindex in $\hat{Q}$. Refer to Triplett (1971), and Kain and Quigley (1975), for discussion of the constant.

(14) $$\hat{A}_j = R_j - \Sigma Z_{ij}\hat{\gamma}_i. \qquad \text{(linear form)}$$

The relationship between the housing index, $\hat{Q}_j$, and the adjusted rent index, $\hat{A}_j$, is given by

(15) $$\hat{Q}_j = \hat{\beta}_0 + \Sigma X_{ij}\hat{\beta}_i$$

$$\hat{A}_j = R_j - \Sigma Z_{ij}\hat{\gamma}_i$$

$$\hat{A}_j = \hat{Q}_j + \hat{\varepsilon}_j.$$

Thus the values of $\hat{Q}_j$ and $\hat{A}_j$ differ by terms involving the error of estimate for a given household.

To the extent that the error term in Equation (10) consists primarily of omitted housing attributes, $\hat{A}_j$ may be preferable to $\hat{Q}_j$ as an index of housing services. To the extent that the error term primarily reflects non-housing factors, $\hat{Q}_j$ is preferable. In addition, $\hat{A}_j$, as compared with $\hat{Q}_j$, has at least two drawbacks. $\hat{A}_j$ will tend to include inflationary factors in the quality index (unless some explicit attempt is made to remove such effects) and to be a less appropriate estimate of quality change for many specific analytical applications. Each of these issues is discussed below.

The $\hat{Q}_j$ index corrects for inflation by taking the index weights from a single time period. In theory, weights may be taken either from the baseline period or from a later (comparison) time period (see Appendix I). In practice, the baseline weights may be better estimates, because the available sample at baseline (enrolled households) is much larger than at subsequent periods (Control households).[1,2]

---

[1] It is inappropriate to use Experimental households other than Control households for estimation in post-enrollment periods, since the experimental treatments may have altered their choice of housing.

[2] The same issues are involved in comparing levels of housing quality at the two sites--Phoenix and Pittsburgh--or in comparing racial submarkets. In each case, households are restricted to markets that may have very different hedonic weight structures. Again, the best solution is to select a base set of weights (see Chapter 4).

Among the major analytical issues to be addressed with hedonic quality measures are the change in quality for households that move and for households that upgrade.[1] In addition, an attempt will be made to determine whether any quality change occurs for households that neither move nor upgrade but whose rent increases; for example, Minimum Rent households that meet the Minimum Rent requirement without moving. The adjusted rent index in effect presumes that any increase in rent for nonmovers reflects changes in quality. The housing index must be used if this presumption is to be tested. However, the housing index will be misleading to the extent that quality changes occur in non-included variables (the error term in Equation (10) includes the effects of omitted quality variables). Nevertheless, the housing index is preferable for most analyses because it is directly affected by changes in quality attributes and is not subject to variation from factors not related to quality. This is crucial for analyses that address differences in quality for nonmovers and differences in shopping ability or incentives. Adjusted rent, on the other hand, may be biased precisely because of omitted variables that relate to tenure conditions, shopping practices, or random price/quality effects and cannot be used to address these issues.

Fortunately, additional information exists which may shed some light on quality variables omitted from the housing index. For households that do not move, survey information on landlord repairs and maintenance may be useful. For households that upgrade to meet Minimum Standards the actual change in the Minimum Standards components, which is not necessarily reflected in the quality index, may be used directly. An increase in satisfaction with the dwelling unit, in conjunction with evidence of repairs, improved maintenance, or Minimum Standards upgrading, may confirm that improvement in "omitted" variables has occurred. For households that move, additional types of quality information may also be relevant. For example, survey data permits analysis of preference achievement and of changes in satisfaction with both the dwelling unit and the neighborhood.

[1] For households in the Housing Gap Minimum Standards treatment, the term upgrade has a specific definition: units that do not meet the Minimum Program Standards may be upgraded to these standards (see Appendix IV).

Also, the move could increase the household's location, in terms of accessibility to workplace or to other facilities.

Hedonic indices can also be used to develop more explicit analyses of factors that affect response to housing allowances, such as shopping ability, racial price discrimination, and tenure conditions. The larger the value of the ratio $\hat{Q}/R$, the greater the quality obtained for a given expenditure. A related issue is whether, for a given change in rent, the change in quality is the same for all Experimental households--that is, whether the relative value of $\Delta\hat{Q}/\Delta R$ differs among experimental groups. Differences in shopping ability, market segmentation, and experimentally induced shopping behavior may all lead to variation in these ratios. The shopping ability of elderly households, for example, may be less than that of younger households. Or, if minority households pay a premium to obtain housing equivalent to that of nonminority households, then the value of $\hat{Q}/R$ will be lower for minorities than for white households. If housing allowance recipients shop less carefully than nonrecipients, the difference between actual rent and the hedonic estimate of market value should be greater for Experimental households than for Control households. Also, it is useful to assess whether households that are required to meet the program Minimum Standards shop more carefully than Minimum Rent or Percent of Rent households in obtaining the most quality for their money. For all these purposes the housing index, $\hat{Q}$, provides estimates, whereas the adjusted rent index, $\hat{A}_j$, does not.

With respect to tenure conditions, some households may obtain a "better deal" than others because of long-term residency or because of a particular relationship with their landlord. Households for which the value of $\hat{Q}/R$ is relatively high may have less incentive to participate in the program; they may be more satisfied with their present units or may have less incentive to move. Estimates of the cost of moving in terms of loss of the long-tenure premium can be used in modeling residential mobility. Estimates of the value of these tenure conditions are obtained as part of the set of hedonic weights, even though they are not used directly in the housing index, $\hat{Q}$.

Finally, hedonic indices estimated for a common set of variables can be used to construct housing price indices for Pittsburgh and Phoenix. This can be important in developing strategies to vary the level of allowance payments across cities.[1] For this purpose the housing index can include the specific items of interest.

All of the above considerations will guide the application of hedonic quality measures in Demand Experiment analyses. Generally, the housing index approach will be used.

---

[1] This has been a major issue in the Section 8 Existing Housing Program concerning the appropriate level of Fair Market Rents.

REFERENCES

Abt Associates Inc., Experimental Design and Analysis Plan, Cambridge, Mass., March 1973, revised August 1973.

Bailey, Martin, "Effects of Race and Other Demographic Factors on the Value of Single-family Homes," Land Economics, vol. 42, May 1966, pp. 215-20.

Berry, Brian and Robert Bednary, "A Hedonic Model of Prices and Assessments for Single-family Homes," Land Economics, February 1975.

Box, G.E. and Cox, D.R., "An Analysis of Transformations," Journal of the Royal Statistical Society, vol. 26, no. 2, 1964, pp. 211-252.

Cheng, David and Harvey J. Iglarsh, "Principal Component Estimators in Regression Analysis," Review of Economics and Statistics, vol. 58, no. 2, May 1976, pp. 229-234.

Daniels, Charles, "The Influence of Racial Segregation on Housing Prices," Journal of Urban Economics, vol. 2, 1975, pp. 105-22.

Edwards, John B., "The Relation Between the F-test and $\bar{R}^2$," The American Statistician, vol. 23, December 1969, p. 28.

Friedman, Joseph and Stephen D. Kennedy, Housing Expenditures Under a Housing Gap Housing Allowance, Cambridge, Mass., Abt Associates Inc., May 1977.

Gillingham, Robert, "Place-to-Place Rent Comparisons Using Hedonic Quality Adjustment Techniques," Research Paper No. 7, Washington, D.C., Bureau of Labor Statistics, March 1973.

________________, "Place-to-Place Rent Comparisons Using Hedonic Quality Adjustment Techniques," Bureau of Labor Statistics Staff Paper No. 8, Washington, D.C., Bureau of Labor Statistics, 1975.

Glejser, H., "A New Test for Heteroskedasticity," Journal of the American Statistical Association, vol. 64, 1969, pp. 316-23.

Grether, D.M. and Peter Mieszkowski, "Determinants of Real Estate Values," Journal of Urban Economics, vol. 1, no. 2, April 1974, pp. 127-146.

Haitovsky, Yoel, "A Note on the Maximization of $\bar{R}^2$," The American Statistician, vol. 23, no. 1, February 1969, pp. 20-21.

Haugens, Robert and A. James Heins, "A Market Separation Theory of Rent Differentials," Quarterly Journal of Economics, vol. 83, November 1969, pp. 660-72.

Hocking, R.R., "Misspecification in Regression," The American Statistician, vol. 28, no. 1, February 1974, pp. 39-40.

Kain, John and John Quigley, "Measuring the Value of Housing Quality," Journal of the American Statistical Association, vol. 65, June 1970, pp. 532-48.

____________, Housing Markets and Racial Discrimination, New York, National Bureau of Economic Research, 1975.

King, Thomas and Peter Mieszkowski, "Racial Discrimination, Segregation, and the Price of Housing," Journal of Political Economy, vol. 81, May-June 1973, pp. 590-606.

King, Thomas, "The Demand for Housing: A Lancastrian Approach," (paper prepared for the Conference on Research in Income and Wealth, November 30-December 7, 1973), New York, National Bureau of Economic Research, 1973.

____________, Property Taxes, Amenities, and Residential Land Values, Cambridge, Mass., Ballinger Publishing Co., 1973.

Lapham, Victoria, "Do Blacks Pay More for Housing?", Journal of Political Economy, vol. 79, November-December, 1971, pp. 1244-57.

Marquandt, D. and Snee, R., "Ridge Regression in Practice," The American Statistician, vol. 29, 1975, pp. 3-19.

Merrill, Sally Roe, The Effects of Segregation and Discrimination on the Price and Quality of Minority Housing, Ph.D Dissertation, Boston University, 1976.

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

Quigley, John, "Housing Demand in the Short Run; An Analysis of Polytomous Choice," (paper prepared for the Winter Meetings of the Econometric Society, December 1973, New York), Yale University Institute for Social and Policy Studies, 1973.

____________, "Racial Discrimination and the Housing Consumption of Black Households," Patterns of Racial Discrimination, Vol. I: Housing, edited by George M. vonFurstenberg, Lexington, Mass., D.C. Heath, 1974.

Rao, Potluri, "Some Notes on Misspecification in Multiple Regressions," The American Statistician, vol. 25, December 1971, pp. 37-39.

Rao, Potluri and Roger Miller, Applied Econometrics, Belmont, California, Wadsworth Publishing Co., 1971, pp. 107-111.

Rapkin, Chester, "Price Discrimination Against Negroes in the Rental Housing Market," Essays in Urban Land Economics, Los Angeles and Berkeley, University of California Press, 1966.

Ridker, Ronald and John Henning, "Determinants of Residential Property Values with Special Reference to Air Pollution," Review of Economics and Statistics, May 1967.

Schnare, Ann, Externalities, Segregation, and Housing Prices, Washington, D.C., Urban Institute, 1974.

Schnare, Ann and Raymond Struyk, "Segmentation in Urban Housing Markets," (paper presented at the Committee on Urban Economics, Conference on Housing Economics, October 1974, St. Louis, Missouri, Washington University), Washington, D.C., Urban Institute, 1974.

Snee, R., "Some Aspects of Nonorthogonal Data Analysis: Part I, Developing Prediction Equations," Journal of Quality Technology, vol. 5, 1973, pp. 67,79.

Statistical Package for the Social Sciences, 2nd ed., New York, McGraw-Hill, Inc., 1975.

Stengel, Mitchell, Market Separation and Racial Price Differentials: The Ghetto Housing Market, unpublished, Lansing, Michigan State University, May 1973.

Strazheim, Mahlon, "Housing Market Discrimination and Black Housing Consumption," Quarterly Journal of Economics, 1973, pp. 19-43.

________________, An Econometric Analysis of the Urban Housing Market, New York, National Bureau of Economic Research, 1975.

Triplett, Jack, "The Theory of Hedonic Quality Measurement and Its Use in Price Indexes," Bureau of Urban Statistics Staff Paper No. 6, Washington, D.C., Bureau of Labor Statistics, U.S. Department of Labor, 1971.

CHAPTER 3

RESULTS OF THE HEDONIC EQUATIONS

This chapter presents the final regression specifications of the linear and log linear hedonic equations for Pittsburgh and Phoenix. For both equations, variables for number of rooms and square feet per room are specified in natural log form, and the variable for length of tenure is an exponential function. Thus, the basic estimating equations are:

(1) $$R_j = \beta_0 + \beta_1[1-\exp(X_{1j})] + \sum_{i=2}^{3} \beta_i \ln X_{ij} + \sum_{i=4}^{n} \beta_i X_{ij} + u_j$$

and

(2) $$\ln R_j = \beta_0 + \beta_1[1-\exp(X_{1j})] + \sum_{i=2}^{3} \beta_i \ln X_{ij} + \sum_{i=4}^{n} \beta_i X_{ij} + u_j,$$

where

R_j = the rent for the j^{th} unit

X_{1j} = length of the household's residence in the j^{th} unit

X_{2j}, X_{3j} = number of rooms and square feet per room in the j^{th} unit

$X_{4j} \ldots X_{nj}$ = other characteristics.

The means and standard deviations of the variables included in the equations are listed in Table 3-1.[1] The estimated coefficients for the semilog equations are given in Tables 3-2 and 3-4, and for the linear equations in Tables 3-3 and 3-5. The estimated linear and semilog equations are similar in terms of the signs and relative magnitudes of the coefficients, and discussion of the results will generally refer to both equations.

Section 3.1 gives an overall assessment of the results. A more specific discussion of the results for individual variables and of the approach taken to define these variables is presented in Section 3.2. Section 3.3

[1] Table 3-1 lists the means and standard deviations for all variables used in the final analysis, that is, variables for the linear and semilog full sample equations, for the common site equation, and for the minority submarket equations. A complete list of all the variables tested is given in Appendix VI. The acronyms of the variables have been included in Table 3-1 to facilitate reference to the variable definitions given in the appendix. A small number of the variables in Table 3-1 refer to equations discussed in Chapter 4.

Table 3-1

MEANS AND STANDARD DEVIATIONS

PITTSBURGH

VARIABLE DESCRIPTION	ACRONYM	MEAN	STANDARD DEVIATION
Tenure Characteristics			
Related to landlord (0,1)	XRELATED	071	258
Length of residence (exponential function)	XEXP4	441	.370
Length of residence (natural log)	XLNLING	3.472	1.112
Landlord lives in the building (0,1)	XLLBLG	.098	.297
Number of persons per room	XOCCRM	.696	.334
Number of landlord contacts for maintenance	XCONTACT	1.337	1.382
Dwelling Unit Features:			
Area per room (natural log)	XLAREAPR	4.847	181
Total number of rooms (includes kitchen and bath) (natural log)	XLTOTRMS	1.674	.262
Building age (years)	XAGE	49.987	13.912
Stove and refrigerator provided (0,1)	XSTAREF	.109	.312
Stove or refrigerator provided (0,1)	XSTOREF	.169	.375
Inferior or no heat (0,1)	XBADH	.216	.412
Central heat present (0,1)	XCENH	.528	.499
Garage provided (0,1)	XGAR	.064	.244
Offstreet parking provided (0,1)	XOFFSTR	.086	.281
Overall evaluator rating (4 point scale)	XRATINGR	1.780	.642
Dishwasher and/or disposal provided (0,1)	XAPPL	.054	.227
Recent interior painting or papering (0,1)	XPAINT	.100	.300
Average surface and structural quality (4 point scale)	XQUAL2	2.171	.366
Many high quality features (0,1)	XFANCY2	.040	.197
Poor wall and ceiling surface (factor score)	XF2SUR	.010	1.052
Poor window condition (factor score)	XF4WIN	.008	986
Poor bathroom wall and ceiling surface (factor score)	XF6BSUR	.0003	1.070
Adequate light and ventilation (0,1)	XHLIVER	413	.493
Presence of adequate ceiling height (0,1)	XHCEHTR	908	.288
High quality kitchen (0,1)	XKITCHOK	.081	.273
Presence of adequate exits (0,1)	XHADQEXR	.922	.269
Air-conditioning present (0,1)	XACPITT	.111	314
Large multifamily structure (0,1)	XMULTI5	.139	.346
Adequate kitchen facilities present (0,1)	XKITCHP	.992	.087
Working condition of plumbing (5 point scale)	XPLUMW	3.575	.888
Plumbing present (0,1)	XPLUMP	.887	.317
Adequate plumbing present and working (0,1)	XHPLUMR	.830	.376
Presence of private yard (0,1)	XYARD	.367	.482
Parking facilities provided (0,1)	XPARK	.148	.355
Temperature control central heat or air-conditioning (0,1)	XTEMP	.576	.494
Neighborhood Features:			
Good recreational facilities and access (factor score)	XCNHF11	-.003	.993
Traffic and litter problems (factor score)	XCNHF13	-.0005	.961
Problems with crime and public services (factor score)	XCNHF14	-.011	.935
Census tracts with higher priced units and higher socio-economic status (factor score)	XCENFO2	-.032	.956
Nonminority census tracts with higher socioeconomic status (factor score)	XCENFO3	.022	.981
Blue collar workers and nonminority residents in census tracts (factor score)	XCENFO4	-.015	1.000
Census tracts with higher socioeconomic status (factor score)	XCTFO2	.009	.929
Census tracts with newer, higher priced units (factor score)	XCTFO3	-.032	.892
Median income of census tract (dollars)	XCTMDINC	8502.807	1623.467
Quality of adult recreation facilities	XHCNAREC	1.417	296
High quality block face (0,1)	XFANCYN	.372	.484
Distance from Central Business District (miles)	XDIST	5.480	3.724
Quality block face landscaping (4 point scale)	XLNDSCPR	1.375	.934
Rent.			
Analytic rent	XACRA61H	111.052	32.396
Natural logarithm of analytic rent	XLACR61H	4.667	.293

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 3-1 (continued)

PHOENIX

VARIABLE DESCRIPTION	ACRONYM	MEAN	STANDARD DEVIATION
Tenure Characteristics:			
Related to landlord (0,1)	XRELATED	058	233
Length of residence (exponential function)	XEXP366	.280	319
Length of residence (natural log)	XLNLING	2.743	1.022
Landlord lives in the building (0,1)	XLLBLG	096	.295
Number of persons per room	XOCCRM	.840	.467
Number of landlord contacts for maintenance	XCONTACT	1.269	1.383
Dwelling Unit Features:			
Area per room (natural log)	XLAREAPR	4.688	.199
Total number of rooms (includes kitchen & bath) (natural log)	XLTOTRMS	1.592	.252
Building age (years)	XAGE	24.447	15.170
Stove and refrigerator provided (0,1)	XSTAREF	640	.480
Stove or refrigerator provided (0,1)	XSTOREF	.793	.405
Inferior or no heat (0,1)	XBADH	.314	.464
Central heat present (0,1)	XCENH	.326	.469
Garage provided (0,1)	XGAR	.044	.207
Garage or carport provided (0,1)	XCARGAR	315	.465
Offstreet parking provided (0,1)	XOFFSTR	.512	.500
Overall evaluator rating (4 point scale)	XRATINGR	1.946	.927
Dishwasher and/or disposal provided (0,1)	XAPPL	.159	.366
Recent interior painting or papering (0,1)	XPAINT	.203	.402
Average surface and structural quality (4 point scale)	XQUAL2	2.285	.633
Many high quality features (0,1)	XFANCY2	.126	332
Adequate light and ventilation (0,1)	XHLIVER	.389	.488
Presence of adequate ceiling height (0,1)	XHCEHTR	.906	.292
High quality kitchen (0,1)	XKITCHOK	.212	.409
Presence of adequate exits (0,1)	XHADQEXR	991	.094
Central air-conditioning present (0,1)	XCACPHX	.244	.430
Large multifamily structure (0,1)	XMULTI5	146	.353
Adequate kitchen facilities present (0,1)	XKITCHP	977	.149
Working condition of plumbing (5 point scale)	XPLUMW	3.568	.859
Plumbing present (0,1)	XPLUMP	.920	.271
Adequate plumbing present and working (0,1)	XHPLUMR	.838	.369
Presence of private yard (0,1)	XYARD	522	.500
Parking facilities provided (0,1)	XPARK	.312	.463
Temperature control: central heat or air-conditioning (0,1)	XTEMP	344	.475
Neighborhood Features:			
Overall neighborhood quality (factor score)	XCNHF11	002	1.002
Recreational facilities (factor score)	XCNHF12	.024	.987
Access to shopping and parking (factor score)	XCNHF14	.007	.998
Census tracts with higher priced units and higher socio-economic status (factor score)	XCENF01	.012	994
Owner-occupied single-family dwelling units in census tract (factor score)	XCENF02	-.016	.992
Poor quality housing in census tract (factor score)	XCENF03	.002	.973
Census tracts with higher priced units and higher socio-economic status (factor score)	XCTF01	.006	1.060
Owner-occupied single-family dwelling units in census tract (factor score)	XCTF02	-.026	1.031
Poor quality housing in census tract (factor score)	XCTF03	-.006	1.474
Median income of census tract (dollars)	XCTMDINC	8072.137	2148.115
Quality of adult recreation facilities	XHCNAREC	1.597	.274
High quality block face (0,1)	XFANCYN	.504	.500
Distance from Central Business District (miles)	XDIST	5.382	4.290
Quality block face landscaping (4 point scale)	XLNDSCPR	1.697	.821
Rent:			
Analytic rent	XACRA61H	132.544	45.545
Natural logarithm of analytic rent	XLACR61H	4.822	.372

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 3-2

SEMILOG EQUATION PITTSBURGH

	R^2 = 0 662 $\bar{R}^2$ = 0 654 F = 89.140 N = 1,583		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Charac-teristics	Related to landlord (0,1)	-0.102	5 813
	Length of residence (exponential function)	-0.141	11.570
	Landlord lives in the building (0,1)	-0.067	4.376
	Number of persons per room	0.082	5.946
	Number of landlord contacts for maintenance	0.012	3.491
Dwelling Unit Features	Area per room (natural log)	0.170	6.449
	Total number of rooms (natural log)	0 565	29.073
	Building age (years)	-0.002	4.168
	Stove and refrigerator provided (0,1)	0 111	6.382
	Inferior or no heat (0,1)	-0.077	6.403
	Garage provided (0,1)	0.091	4.912
	Offstreet parking provided (0,1)	0.022	1.352
	Overall evaluator rating (4 point scale)	0.053	5.846
	Dishwasher and/or disposal provided (0,1)	0 054	2.692
	Recent interior painting or papering (0,1)	0.052	3 497
	Many high quality features (0,1)	0.038	1 576
	Poor wall and ceiling surface (factor score)	-0.019	4.020
	Poor window condition (factor score)	-0.018	3.697
	Poor bathroom wall and ceiling surface (factor score)	-0.013	2.992
	High quality kitchen (0,1)	0.034	1.982
	Presence of adequate exits (0,1)	0.046	2.709
	Air-conditioning present (0,1)	0.025	1.698
	Presence of adequate ceiling height (0,1)	0.034	2.170
	Adequate kitchen facilities present (0,1)	0.117	2.267
	Large multifamily structure (0,1)	0.038	2.527
	Working condition of plumbing (5 point scale)	0.008	1.539
	Presence of private yard (0,1)	0.015	1.468
Neighborhood Features	Good recreational facilities and access (factor score)	0.024	4 964
	Traffic and litter problems (factor score)	-0.009	1.607
	Problems with crime and public services (factor score)	-0.015	2.926
	Census tracts with higher priced units and higher socioeconomic status	0.032	5.626
	Nonminority census tracts with higher socio-economic status	0.032	5.542
	Blue collar workers and nonminority residents in census tracts	-0.026	5.694
	High quality block face (0,1)	0.043	4.160
CONSTANT		2.629	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic ≥ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table 3-3

LINEAR EQUATION PITTSBURGH

$R^2 = 0.656$ $\bar{R}^2 = 0.648$ $F = 93.135$ $N = 1,599$

	VARIABLE DESCRIPTION	COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-11.945	6.132
	Length of residence (exponential function)	-15.036	11.070
	Landlord lives in the building (0,1)	-5.385	3.153
	Number of persons per room	7.651	4.955
	Number of landlord contacts for maintenance	1 073	2.922
Dwelling Unit Features	Area per room (natural log)	19.708	6.789
	Total number of rooms (natural log)	60 020	28.697
	Building age (years)	-0.232	5.202
	Stove and refrigerator provided (0,1)	14.715	7.658
	Inferior or no heat (0,1)	-6.790	5.097
	Garage provided (0,1)	14.379	7.022
	Offstreet parking provided (0,1)	2.837	1.571
	Overall evaluator rating (4 point scale)	5.170	5.187
	Dishwasher and/or disposal provided (0,1)	9.376	4.146
	Recent interior painting or papering (0,1)	6 292	3.801
	Many high quality features (0,1)	8.916	3.311
	Poor wall and ceiling surface (factor score)	-1.670	3.147
	Poor window condition (factor score)	-2.236	4.114
	Poor bathroom wall and ceiling surface (factor score)	-1.627	3.342
	High quality kitchen (0,1)	5.657	2.927
	Presence of adequate exits (0,1)	4 505	2.366
	Air-conditioning present (0,1)	3.171	1.934
	Presence of adequate ceiling height (0,1)	3.038	1.746
	Adequate kitchen facilities present (0,1)	6.575	1.158
	Large multifamily structure (0,1)	3 292	1.986
Neighborhood Features	Good recreational facilities and access (factor score)	2.496	4.706
	Traffic and litter problem (factor score)	-1.112	1.797
	Problems with crime and public services (factor score)	-1.462	2.570
	Census tracts with higher priced units and higher socioeconomic status	3.677	5.890
	Nonminority census tracts with higher socio-economic status	3.691	5.833
	Blue collar workers and nonminority residents in census tract	-2.722	5.488
	High quality block face (0,1)	5.274	4.643
CONSTANT		-100.782	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table 3-4

SEMILOG EQUATIONS PHOENIX

	R^2 = 0 804	$\bar{R}^2$ = 0 801 F = 238.060	N = 1,593
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Charac-teristics	Related to landlord (0,1)	-0.129	7.037
	Length of residence (exponential function)	-0.195	13 508
	Number of persons per room	0.064	6.287
	Number of landlord contacts for maintenance	0.014	4.463
Dwelling Unit Features	Area per room (natural log)	0.310	13.146
	Total number of rooms (natural log)	0.679	34.543
	Building age (years)	-0.002	5.330
	Stove or refrigerator provided (0,1)	0 032	2,549
	Central heat present (0,1)	0.039	2.744
	Garage or carport provided (0,1)	0.031	3.128
	Dishwasher and/or disposal provided (0,1)	0.036	2.486
	Recent interior painting or papering (0,1)	0.015	1.391
	Average surface and structural quality (4 point scale)	0 125	9 571
	Adequate light and ventilation (0,1)	0.035	3.665
	Central air-conditioning present (0,1)	0.050	3.132
	Large multifamily structure (0,1)	0.023	1.674
	Plumbing present (0,1)	0.046	2.507
	Inferior or no heat (0,1)	-0.026	2.049
	Presence of adequate ceiling height (0,1)	0.020	1.279
Neighborhood Features	Overall neighborhood quality (factor score)	0.019	3.284
	Recreational facilities (factor score)	0.016	3.144
	Access to shopping and parking (factor score)	0.013	2.265
	Census tracts with higher priced units and higher socioeconomic status	0.025	3.266
	Owner-occupied, single-family dwelling units in census tract	0.006	1.025
	Poor quality housing in census tracts	-0.029	5.559
	Distance from Central Business District (miles)	-0.004	3.611
	Quality of block face landscaping (4 point scale)	0.021	3.867
CONSTANT		1.902	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table 3-5

LINEAR EQUATION PHOENIX

R^2 = 0 786 $\bar{R}^2$ = 0.783 F = 240 505 N = 1,593

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Charac-teristics	Related to landlord (0,1)	-15.237	6.544
	Length of residence (exponential function)	-22.758	12.330
	Number of persons per room	7 573	5.871
	Number of landlord contacts for maintenance	1.134	2.887
Dwelling Unit Features	Area per room (natural log)	36.257	12.276
	Total number of rooms (natural log)	79.480	33.024
	Building age (years)	-0.251	4.398
	Stove or refrigerator provided (0,1)	4.338	2.717
	Central heat present (0,1)	8.290	4.650
	Garage or carport provided (0,1)	4.501	3.567
	Dishwasher and/or disposal provided (0,1)	8.750	4.737
	Recent interior painting or papering (0,1)	2.078	1.498
	Average surface and structural quality (4 point scale)	14.298	9.364
	Adequate light and ventilation (0,1)	6.512	5 278
	Central air-conditioning present (0,1)	6.802	3.366
	Large multifamily structure (0,1)	4.195	2.344
Neighborhood Features	Overall neighborhood quality (factor score)	2.294	3.156
	Recreational facilities (factor score)	2.480	3.792
	Access to shopping and parking (factor score)	0.972	1.308
	Census tracts with higher priced units and higher socioeconomic status	3.851	4.024
	Owner-occupied, single-family dwelling units in census tracts	1.567	2.280
	Poor quality housing in census tracts	-2.936	4 469
	Distance from the Central Business District (miles)	-0.530	3.555
	Quality of block face landscaping (4 point scale)	2.681	3.856
CONSTANT		-207 014	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES· Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

discusses the analysis of residuals from the hedonic regressions, which was performed in order to determine whether the formal assumptions of the least squares model are reasonably met and whether the residuals show systematic association with external variables, such as geographic or demographic characteristics. Finally, Section 3.4 discusses adjustments to rent (the dependent variable in the hedonic regression).

3.1 GENERAL RESULTS

Overall, the results of the estimation are very satisfactory. The criteria set forth in Chapter 2 for selecting a "best" equation--explanatory power, a broadly representative group of significant attributes, and a reduction in collinearity among individual variables--appear to be reasonably well met. In Pittsburgh, 66 percent of the variance of the log of rent is explained by the available data and in Phoenix, 80 percent. Corresponding figures for the linear equations are 66 percent in Pittsburgh and 79 percent in Phoenix. These results compare very favorably with the explanatory power obtained in other studies that used individual dwelling unit data.

A large number of attributes expected to be important in describing housing bundles are in fact significant and have the expected sign.[1] The variables represent all major component groups of the housing bundle--tenure conditions, dwelling unit quality, dwelling unit size, neighborhood quality, and accessibility. And, within most component groups, a broadly descriptive set of variables is significant. Dwelling unit descriptors include basic facilities (such as heat or kitchen facilities), additional features (such as air-conditioning or appliances), and the surface and structural quality of walls, ceilings, and floors. Neighborhood quality is described by the immediate neighborhood (the "block" face of the unit), by the housing and the socio-economic characteristics of the Census tract, and, for aggregations of Census tracts, by numerous measures of amenities and public services as perceived by those enrolled in the Demand Experiment.

---

[1]As discussed in Chapter 2, "significant" in this chapter means significant at the 0.25 level for a two-tailed test and at the 0.125 level for a one-tailed test (i.e., a t-statistic greater than 1).

The broadly representative equation should provide a solid basis for deriving measures of housing quality from the hedonic regressions. Undoubtedly, relevant attributes are missing from the equation. Examples of possibly important omitted variables are type of building material, aesthetic features, lot size, and effective tax rates for different jurisdictions; these data are not available. Nevertheless, because a wide range of both dwelling unit and neighborhood attributes is significant, the residual term probably does not represent variables excluded from any one category, and the index should be sensitive to many types of improvements in housing quality.

The very limited extent of multicollinearity in the final equations is important in terms of the reliability and stability of the coefficients. As discussed in Section 2.2, the degree of collinearity in the predictor matrix is assessed primarily by computing the variance inflation factors and the related coefficient of determination, R_j^2. In only one case in the four final equations does the variance inflation factor approach the rule-of-thumb limit of four.[1] In Phoenix, the average interior quality variable is correlated with other dwelling unit descriptors, especially dwelling unit age.[2]

As might be expected, the final hedonic regressions for Pittsburgh and Phoenix are different. Different variables remain in the final equations and the coefficient values of included variables often differ substantially. As indicated by the mean variable values in Table 3-1, the two housing markets are in fact very different. Phoenix residents generally have lived in their units for less time, reflecting higher mobility rates. Phoenix units are much newer (about half the average age of Pittsburgh units) and more often have features associated with newer units, such as a dishwasher, a disposal, or a stove or refrigerator included with the unit. In addition, Phoenix units tend to have fewer rooms, somewhat smaller rooms, somewhat higher average ratings for surface and structural quality, and higher

[1] As the variance inflation factor for a variable approaches four, the square of the multiple correlation of that variable and the other predictors approaches 0.75.

[2] The variance inflation factors are listed in Appendix VIII, together with the determinant of the correlation matrix.

overall evaluator ratings (the variance of these ratings is also higher in Phoenix).[1]

A few of the variables are included in only one of the cities simply because they vary substantially only in that city. For example, very few units in the Pittsburgh sample have carports or central air-conditioning. In many other cases, however, the variables in both cities have similar means and standard deviations. In these cases, the fact that a given variable is significant in only one city is apparently due to different patterns of correlation among the variables in each city.

Finally, many of the variables have been defined on a site-specific basis. Thus, although the same basic data elements significantly affect rent in both cities, the specification of these data differs across sites. For example, as discussed below in Section 3.2, principal component analysis has been used for highly collinear groups of variables such as interior surface and structural quality, Census tract neighborhood descriptors, and participant ratings of their neighborhoods; this analysis has been performed separately for Pittsburgh and Phoenix.

The higher R^2 achieved in Phoenix probably reflects its greater homogeneity. The Pittsburgh housing stock is generally older and is divided into many more well-defined neighborhoods than the Phoenix housing stock. Variables such as age of building or Census tract characteristics are probably less likely to capture the individual characteristics of units or neighborhoods in Pittsburgh than they are in Phoenix.[2] In Pittsburgh, Census tract variables will more often represent averages of disparate neighborhoods, and buildings of a given age vary more depending on the extent of past maintenance and rehabilitation.

3.2 ATTRIBUTES OF THE HOUSING BUNDLE

Housing attributes are discussed below in terms of four major component groups: tenure characteristics; dwelling unit attributes; neighborhood

---

[1] These differences are tested formally in Chapter 4, which discusses a common equation for Pittsburgh and Phoenix.

[2] In Phoenix, building age probably bears a more distinct relationship to structural features and type of building than in Pittsburgh.

amenities and public services; and accessibility. The discussion that follows is based on Tables 3-2 to 3-5, the semilog and linear equations for Pittsburgh and Phoenix. For ease of reference, the relevant portions of each table are reproduced in the subsections that follow.

Tenure Characteristics

Tenure characteristics are an important source of variation in rent level. Since tenure factors are excluded from the housing index, accurate estimation of their influence is important. Also, tenure characteristics, particularly length of residence, may be important causal variables in analyses of mobility or satisfaction. Tenure characteristics that may be expected to influence rent level include length of residence in the unit, presence of a resident landlord, family relationship to the landlord, and the cost of maintenance in the unit (represented by variables for the number of persons per room and the number of tenant-landlord contacts for maintenance). The coefficients for these variables are shown in Table 3-6 below.

Length of residence. The length of residence is expected to have an inverse relationship with rent. Lease provisions or long-term residence may tend to slow the adjustment of rents to inflation or other changing market conditions. Long-term tenant-landlord relationships may also bring nonmonetary benefits to the landlord or may actually lower the cost of providing housing services. Over long periods, landlords are likely to gain real cost savings from not having to advertise, from not losing rent during temporary vacancies, and possibly from lower maintenance expenditures. On the other hand, tenants with especially good rent deals may be expected to remain longer in their units.

While these factors suggest that a significant discount might be associated with long tenure, the shape of the function is not known. It is unclear, for example, whether the discount begins immediately or after several years; whether it reaches a maximum or continues to increase; and whether it increases at an increasing, decreasing, or constant rate. In order to assess these factors, many functional forms for length of tenure were tested in the final equations. These included several series of dummy variables, as well

Table 3-6

ESTIMATED COEFFICIENTS FOR TENURE VARIABLES

TENURE VARIABLE	LINEAR COEFFICIENT	LINEAR t-STATISTIC	SEMILOG COEFFICIENT	SEMILOG t-STATISTIC
PITTSBURGH				
Related to landlord (0,1)	-11.945	6.132	-.102	5.813
Length of residence (exponential function)	-15.036	11.070	-.141	11.57
Landlord lives in the building (0,1)	-5.385	3.153	-.067	4.376
Number of persons per room	7.651	4.955	.082	5.946
Number of landlord contacts for maintenance	1.073	2.922	.012	3.491
PHOENIX				
Related to landlord (0,1)	-15.237	6.544	-.129	7.037
Length of residence (exponential function)	-22.758	12.330	-.195	13.508
Number of persons per room	7.573	5.871	.064	6.287
Number of landlord contacts for maintenance	1.134	2.887	.014	4.463

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

as continuous variables in linear, logarithmic, exponential, quadratic, and square root form. Most of the forms provided reasonable explanatory power, and no clear-cut choice among them was indicated empirically. A continuous function was preferred since it is more convenient in analysis and is probably less sample-dependent than a long series of tenure dummies. The final variable chosen is a negative exponential function in the form:

$$(3) \qquad D(\text{discount}) = \begin{cases} 0 \text{ for } t < c \\ \beta[1-\exp\ (\alpha(t - c))] \text{ for } t > c, \end{cases}$$

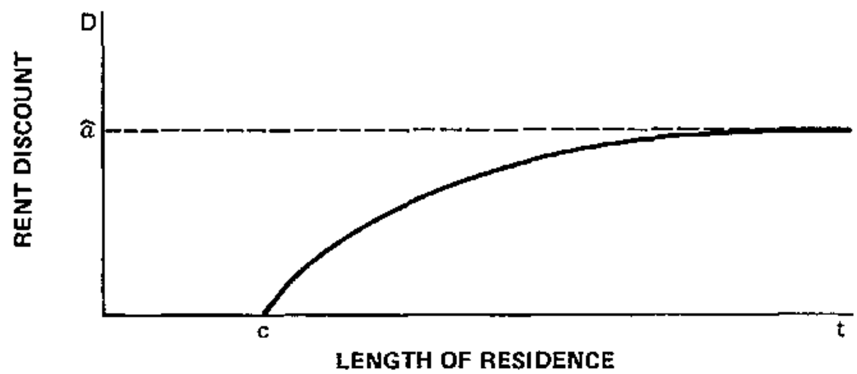
where

t = length of residence

c = length of residence up to which the discount function is zero.

A graph of this form is shown in Figure 3-1.

Figure 3-1
FUNCTIONAL FORM FOR LENGTH OF RESIDENCE VARIABLE



Two types of information were evaluated in deriving this functional form: the change in mean rent for units occupied an additional year, and the regression coefficients of dummy variables that represent each additional year of residence. Graphs of these data indicated that the discount did not begin until after at least 6 months of residence in Phoenix, and 12 months in Pittsburgh. Also, in both cities a maximum discount appears to be reached after about ten years. No attempt was made to estimate the D function directly, since it is nonlinear in its parameters. Since the function is a convenient approximation, it was fitted to the data by hand.

Table 3-7 shows the estimated discounts for various lengths of residence in Pittsburgh and Phoenix. Note that the discount in Phoenix exceeds that in Pittsburgh and takes effect earlier, despite the higher mobility in Phoenix. This difference may simply reflect greater past rent increases in Phoenix or a generally looser market in which landlords increase rents relatively slowly in order to induce tenants to remain.

Other tenure characteristics. When the landlord resides in the dwelling unit, the tenant selection process may be more personally tailored to the landlord's needs, which in turn may lower the cost of providing services. In Pittsburgh, the presence of the landlord in the building results in an average rent discount of about 7 percent as estimated in the semilog equation, and \$5.40 as estimated in the linear equation. This variable is insignificant in Phoenix.

A number of participants stated, in response to an interview question, that they were related to the landlord; these tenants may have paid lower rents than they would have in other comparable units. Indeed, in both cities, a substantial discount appears to be associated with being related to the landlord: \$12.00 or 10 percent in Pittsburgh, and \$15.00 or 13 percent in Phoenix.

Two variables serve as proxies for maintenance costs in a dwelling unit: persons per room and number of requests to the landlord for some item of maintenance. A greater number of people living in the unit, particularly if it is overcrowded, may increase the maintenance costs of the unit. Ideally, maintenance costs for tenants of different characteristics should be examined directly. However, since these data are difficult to collect, the number of persons per room was used as a proxy variable. As seen in Table 3-6, units that are relatively more crowded cost more than equivalent less crowded units in both Pittsburgh and Phoenix. In Pittsburgh, for example, households with one person per room pay about 2 percent more rent than do households with 0.7 persons per room, the sample mean for this variable in Pittsburgh. Although this result presumably reflects increased

Table 3-7

ESTIMATED TENURE DISCOUNTS
(dollars per month)

YEARS OF RESIDENCE	PITTSBURGH		PHOENIX	
	LINEAR	SEMILOG[a]	LINEAR	SEMILOG[a]
1/2	0	0	0	0
1	0	0	3.51	3.67
2	4.26	4.21	8.97	9.21
3	7.32	7.10	12.86	12.94
4	9.50	9.06	15.68	15.58
5	11.07	10.51	17.68	17.52
10	14.28	13.35	21.80	21.19
20	15.01	14.00	22.71	22.02
25+	15.04	14.00	22.76	22.02
Regression Coefficients	$\hat{\beta}$ = 15.036	$\hat{\beta}$ = .141	$\hat{\beta}$ = 22.758	$\hat{\beta}$ = .195

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

NOTE: Refer to Appendix VI for the precise functional form of the tenure variable.

a. Evaluated at the mean rent of 4.6674 in Pittsburgh and 4.8224 in Phoenix.

maintenance costs, other effects may be present as well; persons per room is not an ideal proxy for maintenance costs. To the extent that households give up interior space in order to obtain other dwelling unit or neighborhood attributes not included in the regression, the coefficient of persons per room will tend to reflect these effects as well.

Finally, households were asked in the periodic interviews to indicate how many maintenance items in a prespecified list they had requested from their landlord. Again, while other effects could be present, the moderate increase in rent as the number of requests rose is presumed to represent higher maintenance costs for the unit.

Dwelling Unit Attributes

Dwelling unit variables include basic facilities, higher-quality facilities, safety features, dwelling unit size, and the surface and structural quality of walls, ceilings, and floors. The coefficients and standard errors for these variables are listed in Table 3-8. The major concerns in specifying an appropriate set of dwelling unit variables were to reduce severe collinearity among some of the variables and to explore interactions among attributes whose joint presence might have an independent effect on rent.

Several of the variables used to describe the basic facilities of the dwelling unit represent the components of the Minimum Standards requirement as defined in the Demand Experiment. Six of the 15 components--plumbing facilities, electrical facilities, kitchen facilities, adequate exits, ceiling height, and light and ventilation--were tested in preliminary equations.[1] In addition, the variables that represent plumbing facilities and kitchen facilities were modified to distinguish between the presence of adequate facilities and their working condition.

In the final equations for Pittsburgh, the variables that represent adequate ceiling height and adequate exits are significant and have been retained in both the linear and log linear equations. The variable that indicates the

[1]It should be noted that many other attributes of the unit included in the Minimum Standards are also represented in the equations in a different form. For example, surface and structure ratings for rooms and floors were used as four-point scales rather than as binary (pass/fail) variables.

Table 3-8

COEFFICIENTS OF DWELLING UNIT ATTRIBUTES

DWELLING UNIT VARIABLE	LINEAR COEFFICIENT	LINEAR t-STATISTIC	SEMILOG COEFFICIENT	SEMILOG t-STATISTIC
	PITTSBURGH			
Area per room (natural log)	19.708	6.789	.170	6.449
Total number of rooms (natural log)	60.020	28.697	.565	29.073
Building age (years)	-0.232	5 202	- 002	4 168
Stove and refrigerator provided (0,1)	14.715	7.658	.111	6.382
Inferior or no heat (0,1)	-6 790	5.097	-.077	6.403
Garage provided (0,1)	14.379	7 022	.091	4.912
Offstreet parking provided (0,1)	2.837	1.571	.022	1.352
Overall evaluator rating (4 point scale)	5.170	5.187	.053	5 846
Dishwasher and/or disposal provided (0,1)	9.376	4 146	.054	2.692
Recent interior painting or papering (0,1)	6 292	3 801	.052	3 497
Many high quality features (0,1)	8.916	3.311	.038	1 576
Poor wall and ceiling surface (factor score)	-1 670	3.147	-.019	4.020
Poor window condition (factor score)	-2.236	4 114	- 018	3.697
Poor bathroom wall and ceiling surface (factor score)	-1 627	3.342	-.013	2.992
High quality kitchen (0,1)	5 657	2.927	.034	1.982
Presence of adequate exits (0,1)	4.505	2.366	046	2.709
Air-conditioning present (0,1)	3.171	1.934	025	1.698
Presence of adequate ceiling height (0,1)	3.038	1.746	.034	2.170
Adequate kitchen facilities present (0,1)	6.575	1.158	.117	2.267
Large multifamily structure (0,1)	3.292	1.986	.038	2.527
Working condition of plumbing (5 point scale)	N/A	N/A	.008	1.539
Presence of private yard (0,1)	N/A	N/A	.015	1.468
	PHOENIX			
Area per room (natural log)	36.257	12.276	.310	13.146
Total number of rooms (natural log)	79.480	33 024	679	34.543
Building age (years)	-0.251	4.398	- 002	5.330
Stove or refrigerator provided (0,1)	4 338	2.717	.032	2.549
Central heat present (0,1)	8.290	4.650	.039	2.744
Garage or carport provided (0,1)	4.501	3.567	031	3.128
Dishwasher and/or disposal provided (0,1)	8.750	4.737	.036	2.486
Recent interior painting or papering (0,1)	2.078	1.498	.015	1.391
Average surface and structural quality (4 point scale)	14.298	9 364	125	9.571
Adequate light and ventilation (0,1)	6.512	5.278	035	3.665
Central air-conditioning present (0,1)	6.802	3.366	.050	3.132
Large multifamily structure (0,1)	4.195	2.344	.023	1.674
Plumbing present (0,1)	N/A	N/A	.046	2.507
Inferior or no heat (0,1)	N/A	N/A	-.026	2 049
Presence of adequate ceiling height (0,1)	N/A	N/A	.020	1.279

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

presence of adequate kitchen facilities is also significant in both Pittsburgh equations, whereas the variable that indicates the working condition of the plumbing facilities is included only in the semilog equation. In Phoenix, the presence of adequate light and ventilation is retained in both linear and log linear equations. Two additional basic facilities variables are significant in the semilog equations for Phoenix: adequate ceiling height and the presence of adequate plumbing facilities.

One basic feature of the dwelling unit which was expected to be important, but whose coefficient was less than the standard error, was adequate electrical facilities. This variable is probably misspecified. The currently available data rated units according to a specified number of electrical outlets per room, a factor that probably fails to distinguish units with poor electrical equipment or dangerous wiring conditions from units with satisfactory electrical equipment.

The availability of heat and the type of heating equipment are significant in determining rent level. For example, units with no heat or less desirable types of heating equipment (such as room heaters without flues, fireplaces, stoves, and portable electric heaters) rent for about 8 percent less in Pittsburgh than do units with other types of heating equipment; similarly, the linear equation indicates the discount to be \$6.79. In Phoenix, units with central heating command a premium of about 4 percent (or about \$8.00 based on the linear equation) over units without central heating.

Many other types of facilities provided by the landlord and included in rent are highly valued in both cities. Examples are various types of parking facilities, stove, refrigerator, dishwasher, disposal, and central or other types of air-conditioning.

In order to assess whether the simultaneous presence or absence of a number of features had an independent effect on rent, a series of dummy variables were defined to indicate the joint presence of different groups of attributes. These groups of attributes were selected by examining the simple correlations between the variables and by asking the opinions of the housing evaluators concerning the joint occurrence of attribute types. In Pittsburgh, two of these variables are significant--one that indicates the presence of several high quality kitchen facilities (labeled "high quality kitchen"), and one that indicates the joint presence of a high quality bathroom, a high quality kitchen, all the basic plumbing and electrical facilities, and sound surface and structural conditions (labeled "many high quality features"). Units that have "high quality kitchen" show a premium of about $5.50 (or 3 percent in the semilog equation) over other units. Units that have "many high quality features" show a premium of 4 percent (or about $9.00 in the linear equation) over units that do not have all of them. Furthermore, the overall evaluator rating for the unit, although correlated with many other features, appears to capture additional quality attributes in Pittsburgh.

In Phoenix, none of the joint presence variables nor the overall evaluator rating is significant. This probably means that in Phoenix the surface and structural quality variable is quite important in explaining rent variances and is fairly well correlated with many other dwelling unit quality features. Also, the variance of the "average surface and structural quality" variable is much larger in Phoenix than in Pittsburgh.

Variables that describe the surface and structural quality of the dwelling unit proved somewhat difficult to define. The Housing Evaluation Form provides separate four-point ratings to indicate the surface condition and structural soundness of the walls, ceiling, and floor for each room and of the exterior walls and roof, as well as ratings of the window condition in each room. These ratings are highly correlated, and several approaches to data reduction have been tried. One approach is to average the ratings across rooms, providing separate ratings for the average surface and structural quality of ceilings, walls, and floors, and for average window quality. But these averages also proved to be highly correlated, and the estimated coefficients are very unstable. A second solution is simply to take the average of all the four-point quality ratings for all rooms, the windows, and the exterior. A third approach is to make use of principal components analysis of the individual ratings. Appendix VII presents a discussion of the principal components approach and a listing of the rotated matrices of standardized components.

In Pittsburgh, this third approach has proved quite successful. Six major components have been identified. Analysis of graphic presentations of the rotated standardized components clearly indicates separate clusters of variables. Three of the components are significant in the Pittsburgh equation and have the expected sign. (Since the ratings range from zero ("like new") to three ("needs replacement"), the expected sign of the coefficients on these variables is negative.) The three significant components represent ceiling and wall surface (kitchen, living room, and bedroom); window condition (all rooms); and bathroom ceiling and wall surface. Since the adjusted $\bar{R}^2$ of an equation that uses these variables is higher than the adjusted $\bar{R}^2$ for an otherwise similar equation that uses the average of all quality ratings, the components have been used in the final equation.

In Phoenix, three principal components have been identified. However, the first component was strongly associated with every variable except bathroom and kitchen window condition. The overall average quality variable (labeled "average surface and structural quality") was considered the best representation of surface and structure in Phoenix, since it contributed more to explanatory power than did the variables derived from component analysis. As seen in Table 3-8, a change in surface and structural quality has a very significant impact on rent in Phoenix. A one-unit improvement in the overall average quality rating implies a rent increase of 12.5 percent.

Dwelling unit interior space is described both by the natural log of the total number of rooms and by the natural log of square feet per room. As expected, the relationship between rent and dwelling unit size is nonlinear --rent increases at a decreasing rate with the addition of extra rooms. This is reflected in the logarithmic form of the variable, which is empirically preferable to the linear form. In addition, the variance in the number of square feet per room is quite large and has a very significant influence on rent in both cities.

Finally, dwelling unit type was represented by dummy variables that indicate single-family detached units; single-family attached, row, and duplex units; three- and four-unit structures; multifamily structures with five or more units; and mobile homes. In both Pittsburgh and Phoenix, large multifamily structures are more highly valued than the alternative structure types; large multifamily units command a rent premium of about 4 percent in Pittsburgh and 2 percent in Phoenix, according to the semilog equations.

Neighborhood Amenities and Public Services

The definition of meaningful neighborhood descriptors is complex; neighborhood is both a social and spatial concept. It is difficult to measure public services or neighborhood "status," and difficult to determine the appropriate "spatial" dimension of neighborhood. A spatial concept of neighborhood is clearly relative and differs among individuals. Nevertheless, since most data are available at the block or Census tract level, these geographic areas are most commonly used in practice.

Two studies in particular have attempted to solve some of the problems in defining neighborhood attributes. Kain and Quigley's analysis of the quality of residential environment used 39 separate evaluator ratings on aspects of the interior and exterior of the dwelling unit, the quality of surrounding structures, and the quality of the immediate block face on both sides of the street (Kain and Quigley, 1970a and 1970b). Factor analysis was used to reduce these ratings to five separate dimensions of dwelling unit and neighborhood quality; the derived factors provided descriptions of dwelling unit quality, basic residential quality, and quality of proximate properties, nonresidential land use, and the average structural quality of properties on the block face. Kain and Quigley used these factors as variables in an equation that also included Census variables to reflect a broader neighborhood, expecting that while consumers place great value on the quality of their immediate neighborhood, they are also concerned with a broader area.

Thomas King's study of New Haven examines the use of "subjective" resident opinions, as contrasted with "objective" data, to represent neighborhood and public service quality (King, 1973). A survey of homebuyers provided ordinal ratings (on a 1 to 5 scale) of 13 types of amenities and services, such as school quality, level of crime, street maintenance, and air pollution. These ratings were then averaged over elementary school districts. Although it is difficult to choose an appropriate area for which to average these consumer perceptions, the average responses of a large number of individuals will approximate better than individual judgments the prevailing market judgments about neighborhood qualities (King, 1973). If a household's high rating of a given neighborhood is not shared by other bidders, it may obtain the location at a lower price than it would have been willing to pay, but this individual "bargain" will not affect the market value of the neighborhood. King tested the association between these subjective ratings and "objective" measures of the same attribute (where objective measures were available) and generally found a high correlation between the two.

Both of these studies have influenced the derivation of the neighborhood quality variables employed here. Three spatial dimensions are used--the immediate block face, the Census tract, and an aggregation of Census tracts into larger neighborhoods. In addition, three types of data are used--housing evaluator ratings (recorded on the Housing Evaluation Form), Census data, and participant opinions (recorded in the Baseline Interview). The discussion that follows is organized around these three spatial dimensions. Table 3-9 lists the neighborhood variables included in the hedonic regressions.

Block face variables. The block face represents the area within 100 yards of the unit in both directions. Ratings on aspects of the proximate block face were obtained from the Housing Evaluation Form, and include street maintenance, street lighting, pedestrian walkways, landscaping, street litter, types of surrounding buildings, and the presence of abandoned buildings or cars. In Phoenix, the quality of block face landscaping proved to have a significant effect on rent. In Pittsburgh, a derived dummy variable labeled "high quality blockface" is significant; this variable specifies that several high ratings occur together (for street maintenance, landscaping, and litter) and that no detrimental features are present.

Census tract variables. At the Census tract level, many descriptors of dwelling unit characteristics and of the residents' demographic characteristics are available. As would be expected, these variables are highly collinear. When a number of the Census variables are included in an equation, only a few of them are significant or have the expected sign. Two approaches to finding the most appropriate use of the Census data have been tried--selection of a subset of individual variables, and principal components analysis.

Table 3-9

COEFFICIENTS OF NEIGHBORHOOD ATTRIBUTES

NEIGHBORHOOD VARIABLE	LINEAR		SEMILOG	
	COEFFICIENT	t-STATISTIC	COEFFICIENT	t-STATISTIC
	PITTSBURGH			
Good recreational facilities (factor score)	2.496	4 706	.024	4.964
Traffic and litter problems (factor score)	-1.112	1.797	-.009	1.607
Problems with crime and public services (factor score)	-1.462	2.570	- 015	2 926
Census tract with higher priced dwelling units and higher socioeconomic status (factor score)	3.677	5.890	.032	5.626
Nonminority census tract with higher socioeconomic status (factor score)	3 691	5.833	.032	5.542
Blue collar workers and nonminority residents in census tract (factor score)	-2.722	5.488	-.026	5.694
High quality block face (0,1)	5.274	4.643	.043	4.160
	PHOENIX			
Overall neighborhood quality (factor score)	2.294	3.156	.019	3 284
Recreational facilities (factor score)	2.480	3.792	.016	3.144
Access to shopping and parking (factor score)	0.972	1.308	.013	2.265
Census tract with higher priced dwelling units and higher socioeconomic status (factor score)	3.851	4.024	.025	3.266
Owner-occupied, single-family dwelling units in census tract (factor score)	1.567	2.280	.006	1.025
Poor quality housing in census tract (factor score)	-2.936	4.469	-.029	5.559
Distance from Central Business District (miles)	-0.530	3.555	-.004	3.611
Quality of block face landscaping (4 point scale)	2.681	3.856	.021	3.867

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

When individual Census variables were used in an equation, emphasis centered on selecting measures of dwelling unit quality (such as the proportion of units that are substandard, or the proportion that have air-conditioning) and on measures of the residents' socioeconomic class (such as the proportion of blue-collar workers). Median tract income behaved as a dominant variable; that is, it was highly correlated with other variables and masked their contribution. Even without median tract income, however, the regression results were fuzzy. In one of the preliminary equations estimated for Pittsburgh, only the proportion of substandard units and the proportion of blue-collar workers were significant and had the expected sign. In Phoenix, only the proportion of blue-collar workers was significant. Since these variables are rather arbitrary indicators of neighborhood quality at the Census tract level, an alternative way of using the data was necessary.

Principal components analysis was based on 20 characteristics of tract properties and residents, and analyses were performed separately by site. No "hypotheses" about underlying traits were maintained; rather, the approach was strictly one of data reduction. The components are retained in the equations according to their level of significance.

In Phoenix, many of the variables are highly associated with the first component, which appears to describe the overall socioeconomic status of the tract (labeled "Census tract with higher priced units and higher socioeconomic status"). Thus income, education, and property value are positively associated with this component; overcrowding, blue-collar residents, and Spanish American residents are negatively associated with it. The second component in Phoenix describes neighborhoods with single-family, owner-occupied, larger units. The third component identifies poor quality housing --units without adequate plumbing, without adequate heat, and with incomplete or shared kitchen facilities (labeled "poor quality housing in the Census tract"). These three components are included in both the linear and semilog equations for Phoenix.

In Pittsburgh, the first component represents neighborhoods with higher quality, owner-occupied, single-family units. However, this variable is not included in either equation, probably because of the rather small variance of some of these variables in a sample of low-income renter households. The second and third components in Pittsburgh are significant in both the linear and semilog equations. The second component represents neighborhoods with higher socioeconomic status and newer, higher priced units. Census tracts with a lower proportion of black households and relatively higher income households are represented in the third component. Finally, a fourth component, which is included in the Pittsburgh linear equation, indicates tracts with blue-collar residents and a relatively low proportion of black residents.

There are potential problems with the inclusion in the tract components of demographic descriptors reflecting the overall income, education, occupation, household size, and race/ethnicity of tract residents and of the median rent and value of units in the tract. These variables are included because they appear to be associated with the quality of the neighborhood in terms of its level of services, general ambiance, and overall locational value.[1] It may be argued, however, that the estimated coefficients of factors including these variables will misstate the value of the neighborhood.

The immediate concern with including the median rent and value of units in the tract is that they are somehow tautological. If, in the extreme case, all units in a tract had the same rent, then (mean) tract rent would perfectly predict the rent of units. In this extreme case, inclusion of mean tract rent as a variable would, of course, make it impossible to identify the value of the underlying attributes.[2] On the other hand, there would no longer be any need to do so. Hedonic indices essentially attempt to estimate the normal market value of units. A general consensus that units in a particular area are worth more is no less valid than a consensus that larger units are worth more.

Obviously, of course, to the extent that the variables included in the hedonic index are "black boxes" which predict market cost without explaining it, they may be less likely to be replicated over time or between cities.[3] Indeed, to

---

[1] There are also special issues associated with racial or ethnic descriptors due to the possibility of price differentials introduced by discrimination, as discussed in Chapter 5.

[2] The problem in this case is not, of course, tract rent, but identifying the tract at all. A set of tract dummy variables would raise the same issues.

[3] In addition, potential issues involving market segmentation can only be effectively answered to the extent that the reasons for neighborhood price differences can be ascertained.

the extent that housing markets are dominated by the existing stock and thus can deviate from underlying production costs for long periods of time, the market value of well-defined attributes may also vary over time if demand changes over time. Fortunately, stability can be directly tested in the Demand Experiment, using Control households as discussed in Chapter 4. The tests indicate that the relative attribute prices of the baseline hedonic model are quite stable over the period of the experiment.

In any case, the median rent and value of units in a tract are not by themselves such good predictors of individual unit rents that they threaten to dominate the equation. While the hedonic equations explain 66 and 80 percent of the variance in rent in Pittsburgh and Phoenix, respectively, regressions of ln rent on median tract rent explain 13 and 32 percent and of the variance in ln rent, and regressions based on median tract value explain 13 and 21 percent, respectively.[1]

A more troublesome issue with inclusion of a general tract rent level variable is actually the same issue that arises from inclusion of tract demographic descriptors--the possibility that the value imputed by the estimated coefficients to the tract refers to certain units in the tract which are systematically different from those that will be occupied by other households. This can arise in the following way. Say that, as might be expected, the demographic characteristics of observed households tend to be correlated with tract characteristics. Thus, for example, tracts with higher median income would tend to be associated with households with higher incomes. But higher-income households tend to spend more on housing, including, possibly, housing with attributes not specifically included in the hedonic regressions. In this case, the estimated coefficient for a factor including tract income would in part reflect omitted characteristics and the greater demand for housing by

---

[1]The estimated regressions are:

	PITTSBURGH		PHOENIX	
	Median Census Tract Rent (ln)	Median Census Tract Unit Value (ln)	Median Census Tract Rent (ln)	Median Census Tract Unit Value (ln)
Coefficient	.005	.00003	.006	.00004
F	246.66	250.06	813.53	454.39
R^2	.13	.13	.32	.21
s.d. error of residual	.277	.277	.304	.328
N	1,579	1,579	1,695	1,695

SAMPLE: All enrolled households at the baseline period, excluding those over-income and those having extreme values of hedonic residuals.

higher-income households.[1]

In fact, this does not appear to be a serious problem. Thus, for example, the neighborhood factors included in the hedonic equation are very poor predictors of individual household income, so that there is little reason to suspect that the estimated coefficients reflect the individual households observed rather than the overall characteristics of tract units.

The results of the regression of household income (natural log) on the census tract factor scores included in the final semi-log hedonic equations are shown below in column (1) of Table 3-10. Although the estimated coefficients are significant in Phoenix, the explanatory power is very low (6 percent of the variance in household income is explained by the census tract factor scores). In Pittsburgh, none of the variance is explained by the census tract factor scores.

In order to help assess the potential bias in the estimated coefficients of the census factor scores, an hedonic regression was estimated which corresponded to the final semi-log equation with the following exceptions--the census tract factors were excluded and household income (natural log) was included. The estimated coefficient of household income is .0603 in Pittsburgh and .0625 in Phoenix (and both coefficients are significant at the .01 level).[2]

The potential bias in the estimated coefficients of the census factor scores in the final hedonic regressions would at most be the hedonic coefficient for income times the coefficients of the census factors regressed on household income. Column (2) in Table 3-10 shows this product (the coefficients in column (1) times .0603 in Pittsburgh and .0625 in Phoenix). This result is then compared with the estimated coefficients for the census factor scores for the final hedonic equations, shown in column (3). The difference, as seen in column (4), is generally only in the third decimal place of the hedonic coefficient.

[1]This problem was suggested to us by C. Lance Barnett.

[2]The statistics pertaining to these hedonic regressions are:

	PITTSBURGH	PHOENIX
$\bar{R}^2$	.6396	.7998
F	86.12	257.00
N	1536	1603
Coefficient of (ln) income	.0603	.0674
t-statistic	4.47	6.337

Table 3-10

POTENTIAL BIAS IN THE ESTIMATED COEFFICIENTS
OF THE CENSUS TRACT FACTOR SCORES

(1)		(2)	(3)	(4)
REGRESSION OF HOUSEHOLD INCOME (ln) ON CENSUS TRACT FACTOR SCORES		ESTIMATED COEFFICIENTS-- COLUMN (1) TIMES COEFFICIENT OF INCOME (ln)[a]	ESTIMATED COEFFICIENTS OF CENSUS TRACT FACTOR SCORES IN HEDONIC EQUATION[b]	COLUMN (3) LESS COLUMN (2)
FACTOR	COEFFICIENT (t)			
PITTSBURGH				
Census tracts with higher priced units and higher socioeconomic status (XCENFO2)	.0179 (1.66)	.0011	.02922	.0281
Nonminority Census tracts with higher socioeconomic status (XCENFO3)	.0130 (1.17)	.0008	.02922	.0284
Census tracts with blue collar workers and nonminority residents (XCENFO4)	-.0011 (.10)	-.0001	-.02458	-.0244
$\bar{R}^2$	.0007			
F	1.367			
N	1536			

(continued)

a. Coefficient of natural log of household income in an hedonic regression that includes (ln) income but excludes all Census tract factor scores. See text.

b. Final semi-log hedonic regressions were estimated that correspond to those presented in Tables 3-2 and 3-4 for Pittsburgh and Phoenix. The estimated coefficients presented here for the Census tract factor scores differ very slightly from those in the final equation due to minor differences in sample (arising from assembling a file including only hedonic variables and household income).

Table 3-10 (continued)

POTENTIAL BIAS IN THE ESTIMATED COEFFICIENTS OF THE CENSUS TRACT FACTOR SCORES

(1)		(2)	(3)	(4)
REGRESSION OF HOUSEHOLD INCOME (ln) ON CENSUS TRACT FACTOR SCORES		ESTIMATED COEFFICIENTS-- COLUMN (1) TIMES COEFFICIENT OF INCOME (ln)[a]	ESTIMATED COEFFICIENTS OF CENSUS TRACT FACTOR SCORES IN HEDONIC EQUATION[b]	COLUMN (3) LESS COLUMN (2)
FACTOR	COEFFICIENT (t)			
PHOENIX				
Census tracts with higher priced units and higher socioeconomic status (XCENF01)	.0787 (6.61)	.0049	.0262	.0213
Census tracts with owner-occupied single-family housing (XCENF02)	.0506 (4.18)	.0032	.008	.0048
Poor quality dwelling units in Census tracts (XCENF03)	-.0822 (6.67)	-.0051	-.0324	-.0273
R^2	.0605			
F	35.41			
N	1603			

a. Coefficient of natural log of household income in an hedonic regression that includes (ln) income but excludes all Census tract factor scores. See text.

b. Final semi-log hedonic regressions were estimated that correspond to those presented in Tables 3-2 and 3-4 for Pittsburgh and Phoenix. The estimated coefficients presented here for the Census tract factor scores differ very slightly from those in the final equation due to minor differences in sample (arising from assembling a file including only hedonic variables and household income).

Participant ratings of public services in larger neighborhoods. Participant surveys on neighborhood conditions and public services were also used to specify neighborhood quality, and included ratings of elements such as the quality of schools, police and fire protection, garbage collection service, street lighting, parking, medical and shopping facilities, and the quality of landscaping. Participants also rated the degree to which certain aspects of their neighborhood were a problem, such as the presence of crime, the use of drugs, traffic congestion, poor street maintenance, and abandoned buildings and junk-filled lots. Each of these responses (on either a three- or a four-point scale) has been averaged by Census tract to represent prevailing perceptions of the tract area. Since the sample of participants in most tracts is extremely small, however, a broader definition of neighborhood was necessary.

Fortunately, larger neighborhoods had already been defined. The base payment levels in the Demand Experiment (the C* schedules) were developed from a weighted average of the cost of modest existing standard housing in each neighborhood of the counties, as estimated by a panel of local experts (realtors, HUD area office staff, and so forth). These C* neighborhoods were defined so that they were reasonably homogeneous with respect to housing costs and types and corresponded as much as possible to local perceptions of distinct neighborhoods. Thirty-four such neighborhoods have been defined for Pittsburgh, 20 for Phoenix. Participant ratings have been aggregated over these areas in one of two ways--as the average value of the three- or four-point scale, or (for some variables) as the proportion of households that judged the attribute to be of "poor" quality. The latter method was used when the major sample variation across C* neighborhoods occurred in the proportion of respondents that ranked an attribute as poor.

When these individual variables were used in an equation, the same problem occurred as with the Census data: many of the variables were highly collinear. When more than a few were included in the same equation, the magnitudes and the signs of the coefficients were extremely unstable. In trial equations estimated for Pittsburgh, two variables--problems with litter and trash, and poor quality recreation facilities--were significant, had the expected sign, and provided the most explanatory power. In Phoenix, three variables provided the best results in preliminary equations--the quality of landscaping, elementary schools, and police protection.

Again, in order to identify separate clusters of variables representing neighborhood quality, the full set of participant ratings in each site was used in a principal components analysis. The scores derived from this analysis were tested in the regression, and several are significant. Since the explanatory power of an equation that uses these scores is slightly higher than the explanatory power of an equation that uses a small subset of the separate ratings, and since the components represent a larger number of individual variables, the derived components have been included in the final equations.

The first component in Phoenix (labeled "overall neighborhood quality") is positively associated with many variables that describe the quality of public services (such as schools, police, fire protection, and garbage collection) and the quality of landscaping, and it is negatively associated with neighborhood "problem" ratings (such as poor street maintenance, litter, drugs, crime, abandoned cars, and junk-filled lots). The second significant component in Phoenix represents good quality recreational facilities for adults, teenagers, and children; the third, accessibility to shopping and parking.

In Pittsburgh, the same components are significant in both the linear and semilog equations. The first Pittsburgh component represents areas with good recreational areas, good transportation, and easy access to shopping, medical facilities, and places of worship. The second component represents problems with noise, heavy traffic, drugs, abandoned buildings, and litter. The third indicates areas that have crime problems and poor public services, such as police, fire protection, and elementary schools.

In summary, variables derived from principal component analysis of Census descriptors and of participant perceptions of their neighborhood have been used to describe neighborhood quality. For both types of data, the results seem preferable to the use of only a limited number of individual variables in a given equation. First, the explanatory power of the equation is increased, although the increase is marginal. More important, since the derived components are functions of numerous descriptors of neighborhood quality, they should be more sensitive to changes in the consumption of neighborhood attributes. But neither method--the arbitrary selection of a limited group of individual variables or the use of a data reduction technique such as principal component analysis--is completely satisfying. For example, the assumption in this principal components analysis of a linear relationship among variables such as Census descriptors can certainly be questioned. In addition, the degree of similarity between the characteristics of an individual household and those of neighboring households may affect consumer valuations of neighborhood type.

<u>Accessibility</u>. Ease of access to various facilities is expected to be associated with higher rent levels through its effect on the price of land, which is thought to reflect the cost of transportation from a given location to centers of economic activity. The cost of transportation, in turn, is a function both of the time costs of travel and of operating costs such as bus fares or automobile expenses (Alonso, 1965 and Muth, 1969). Accessibility to work place, in particular, has been hypothesized to play a major role in location decisions. In such models, the price of housing differs across locations because households bid for residential sites that are close to their work places.

Several variables have been defined for use in the present analysis as surrogates for the effects of the land rent gradient. One approach is based on the traditional monocentric model, which assumes that land prices vary as a function of distance from the Central Business District (CBD). Information on travel time to the CBD is not available; instead, distance to the CBD has been used. Both the linear and logarithmic forms were tested in both cities.

In contrast, multicentered models recognize numerous centers of employment. A generalized employment accessibility variable for Pittsburgh, representing the multicentered approach, was obtained from the National Bureau of Economic Research in Cambridge, Massachusetts. This variable is defined by a standard exponential decay function of travel time, weighted by employment in each of the 132 school districts in the Pittsburgh SMSA. Each school district is composed of several Census tracts; the data were allocated to Census tracts for use in this study. Thus,

(4) $$\text{"Generalized employment accessibility"} = \sum_{j=1}^{n} \frac{E_j}{t_{ij}^{\alpha}}$$

where

E_j = employment in zone j

t_{ij} = travel time from zone i to zone j

α = a positive constant.

In Pittsburgh, neither the generalized accessibility variable nor the single distance measure are significant in either the linear or log linear final equation. There are probably two reasons for this. First, the specifications tested are undoubtedly too simple to represent the complex access patterns in Pittsburgh. Second, the principal components analysis based on the aggregations of Census tracts (C* neighborhoods) tends to distinguish between central city and suburban attributes.

Distance from the CBD is included in both Phoenix equations, however, and seems to provide a reasonably strong access measure. Based on the linear equation, rent for units ten miles from the CBD is $5.30 less than for comparable central city units.

3.3 ANALYSIS OF RESIDUALS

The analysis of regression residuals is important in assessing the results of the hedonic model. The residuals should exhibit certain patterns if the formal assumptions of the model are reasonably correct; that is, the errors are normally distributed with zero mean and constant variance. Nonconstant variance (heteroskedasticity) results in inefficient estimation and, if present, suggests some transformation of the data before estimation (Johnston, 1963). Suspect observations with extremely large residuals (outliers) can also be identified. It may be advisable to remove outliers from the final estimation, depending on how they were generated (Anscombe and Tukey, 1963).

More generally, the assessment of residuals with respect to geographic areas and the demographic characteristics of the households will indicate whether the model yields unbiased predictions. This is especially important in light of the analytical uses of both the housing index and the residuals in demand models and in locational choice models. Strong correlation of the residuals with income or demographic groups might indicate either biases resulting from omitted quality variables or market segmentation based on differences in taste.[1] Similarly, severe bunching of residuals according to a geographic pattern suggests either omitted variables or improper specification of neighborhood effects.

---

[1]A related issue--the analysis of submarkets as a function of racial concentration and central city and suburban differences--is discussed in Chapter 4.

The results of the residuals analysis are very encouraging. For both the linear and semilog equations the residuals appear to meet the necessary distributional assumptions. As discussed below, a very small number of cases were designated as outliers and were excluded from final estimation. These outliers are also excluded from the graphic and other analyses of the residuals, which are designed to assess such concerns as heteroskedasticity and nonlinearity and can be difficult to interpret when extreme values are included in the sample. In fact, the tests are not generally sensitive to a single problem, and the inclusion of outlier observations can produce misleading test results. For example, an outlier might increase the variance of the subgroups to which it belongs, or a test might mistakenly suggest heteroskedasticity (Anscombe and Tukey, 1963; Draper and Smith, 1966).

Predictive Ability and Identification of Outliers

The explanatory power of the hedonic regressions is described by the statistics shown in Table 3-11 and by the distributions of prediction error shown in Table 3-12. (The outliers, described below, are excluded from the sample in both tables; 16 observations are designated as extreme in Pittsburgh, 20 in Phoenix.) As shown in Table 3-11, the standard errors of the linear equations are about $19 in Pittsburgh, and $21 in Phoenix. Based on the absolute value of the linear residuals, the mean percentage differences in predicted and actual rent are 14.3 in Pittsburgh and 13.8 in Phoenix; for the log residuals from the semilog equation, the mean percentage difference in log rent and predicted log rent is under 0.03 in both cities.[1] The cumulative distribution of the prediction error, shown in Table 3-12, indicates the tight fit of the model. For both the linear and semilog equations estimated rent for over 75 percent of the households deviated 20 percent or less from actual rent.

It is also useful to visually assess the fit of the model. Figures 3-2 through 3-5 present plots of predicted versus actual values for the semilog and linear equations. In these figures, the actual value is plotted on the vertical axis, the predicted value on the horizontal axis. The

[1] Computed as $\frac{|R-\hat{R}|}{R}$ for the linear equation and as $\frac{|\ln R-\widehat{\ln R}|}{\ln R}$ for the semilog equation.

Table 3-11

DESCRIPTION OF HEDONIC EQUATIONS AND RESIDUALS

DESCRIPTIVE STATISTIC[a]	PITTSBURGH LINEAR	PITTSBURGH SEMILOG	PHOENIX LINEAR	PHOENIX SEMILOG
Standard error of the regression equation	19.206	.17198	21.211	.16587
R^2	.66	.66	.79	.80
$\bar{R}^2$	.65	.65	.78	.80
Minimum residuals	-67.308	-.596	-65.664	-.612
Maximum Residuals	70.174	.593	78.190	.593
Mean percent difference, predicted and actual values[b]	14.3%	2.9%	13.8%	2.7%
Sample	1,599	1,583	1,593	1,593
Extreme residual sample size	16	16	20	20
Mean value	46.239	.133	63.23	.169
Maximum values	-106.105	-1.069	-77.808	-1.233
Maximum values	186.809	1.134	200.749	1.125

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. The statistics refer to the sample which excludes the extreme residuals.

b. Computed as $100|R-\hat{R}|/R$ for the linear equation and $100|\ln R-\widehat{\ln R}|/\ln R$ for the semilog equation.

Table 3-12

DISTRIBUTION OF PREDICTION ERROR
(Cumulative Frequency)

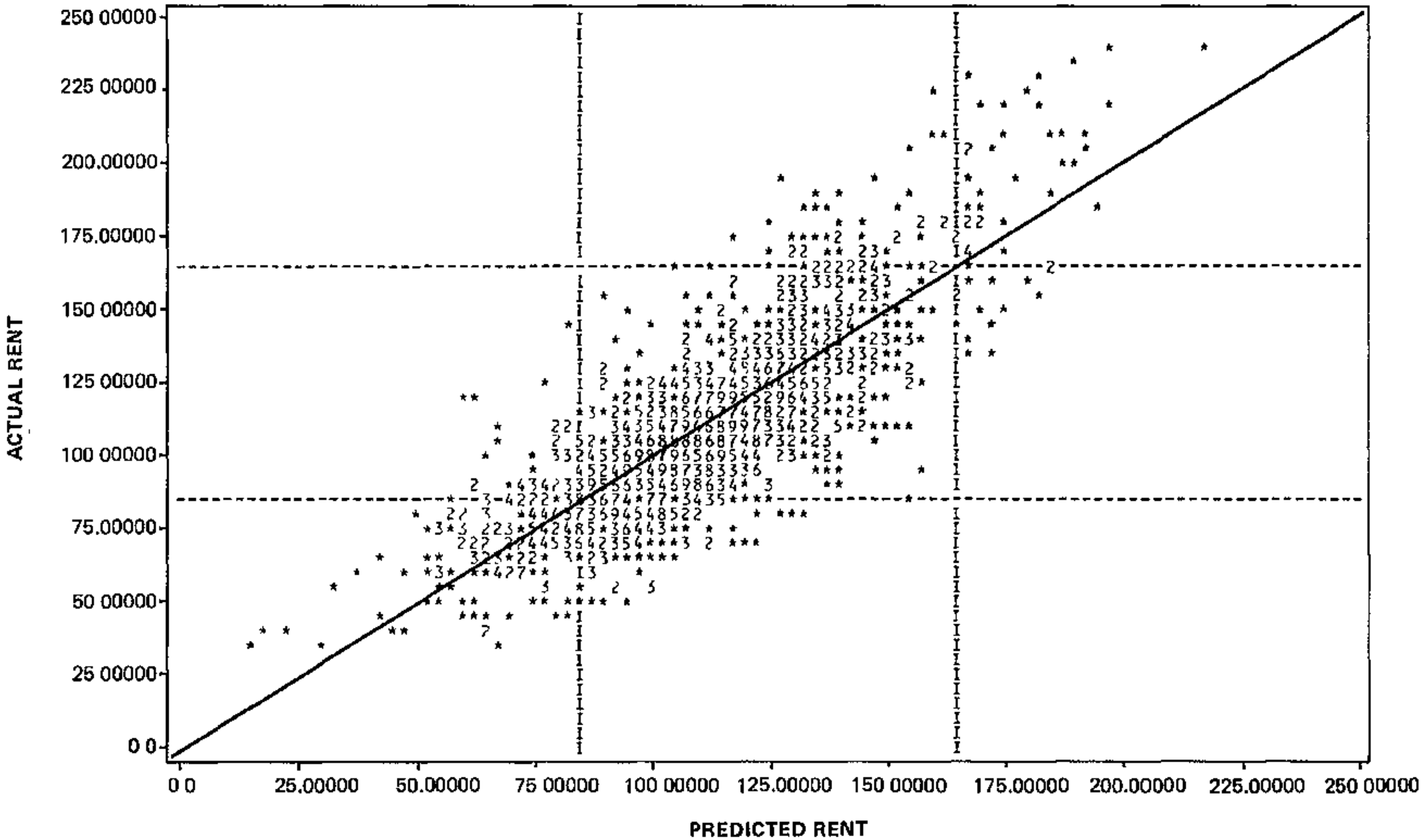
PERCENTAGE DIFFERENCE PREDICTED AND ACTUAL RENT[a]	LINEAR EQUATIONS		SEMILOG EQUATIONS	
	PITTSBURGH	PHOENIX	PITTSBURGH	PHOENIX
0-1	4.8	6.9	5.0	6.4
1-2	9.3	12.1	10.9	11.3
2-3	13.6	18.1	15.9	16.9
3-4	19.3	23.3	20.6	22.4
4-5	24.7	28.5	25.9	27.4
5-10	46.5	50.2	46.7	51.3
10-15	62.9	66.8	65.5	68.4
15-20	76.8	77.7	78.1	79.9
20-25	84.9	85.2	86.7	88.4
25-30	91.0	90.1	92.5	93.1
30-40	95.9	94.9	96.8	97.2
40-50	97.7	97.7	98.1	98.6
50-75	99.6	99.4	99.7	99.7
75-100	100.0	100.0	100.0	100.0
> 100	--	--	--	--
SAMPLE	1,599	1,589	1,583	1,594

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Computed as $|R-\hat{R}|/R$ for the linear equation and $|\exp \ln R - \exp \widehat{\ln R}|/\exp \ln R$ for the semilog equation.

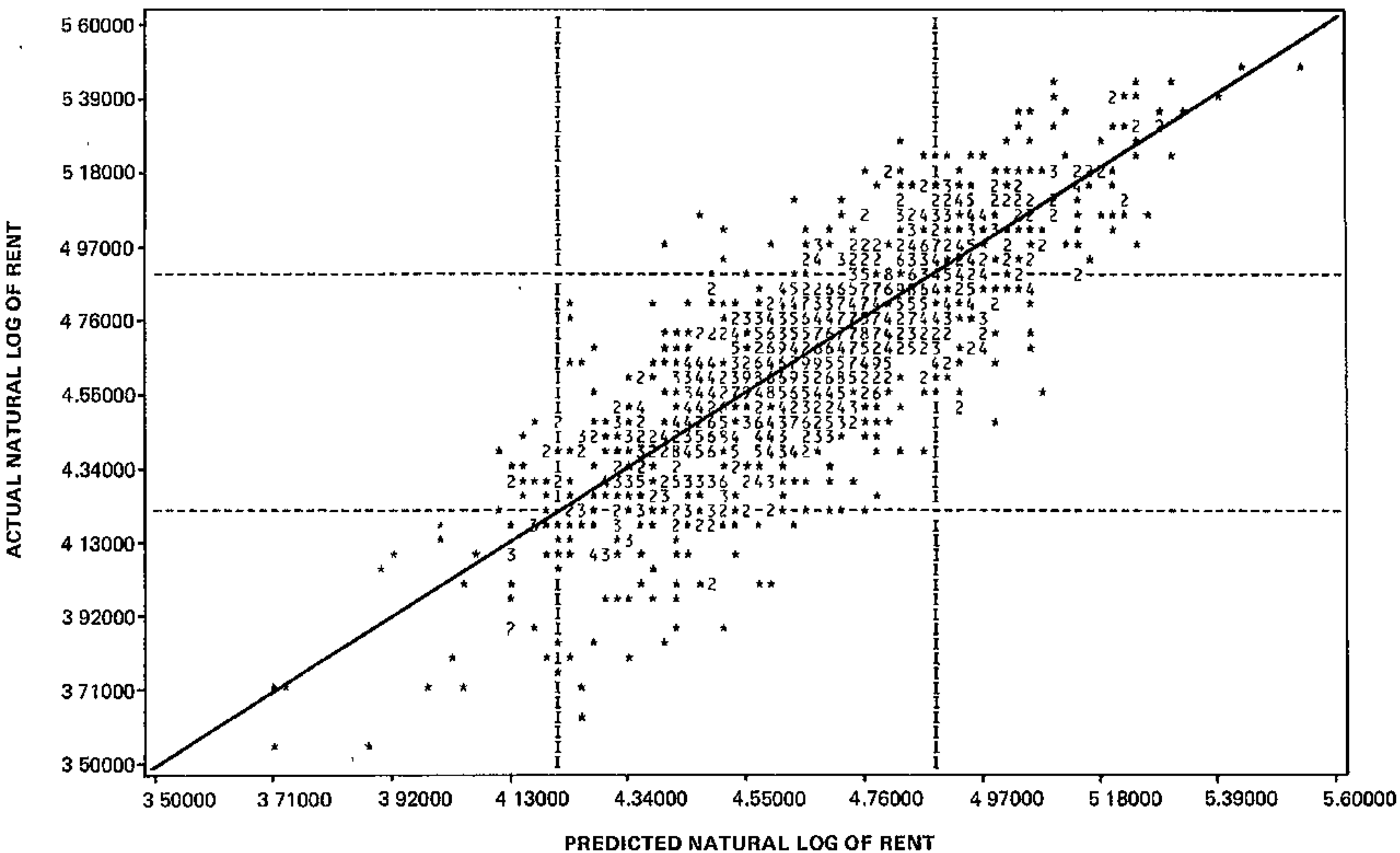
Figure 3-2
ACTUAL AND PREDICTED RENT: PITTSBURGH



SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

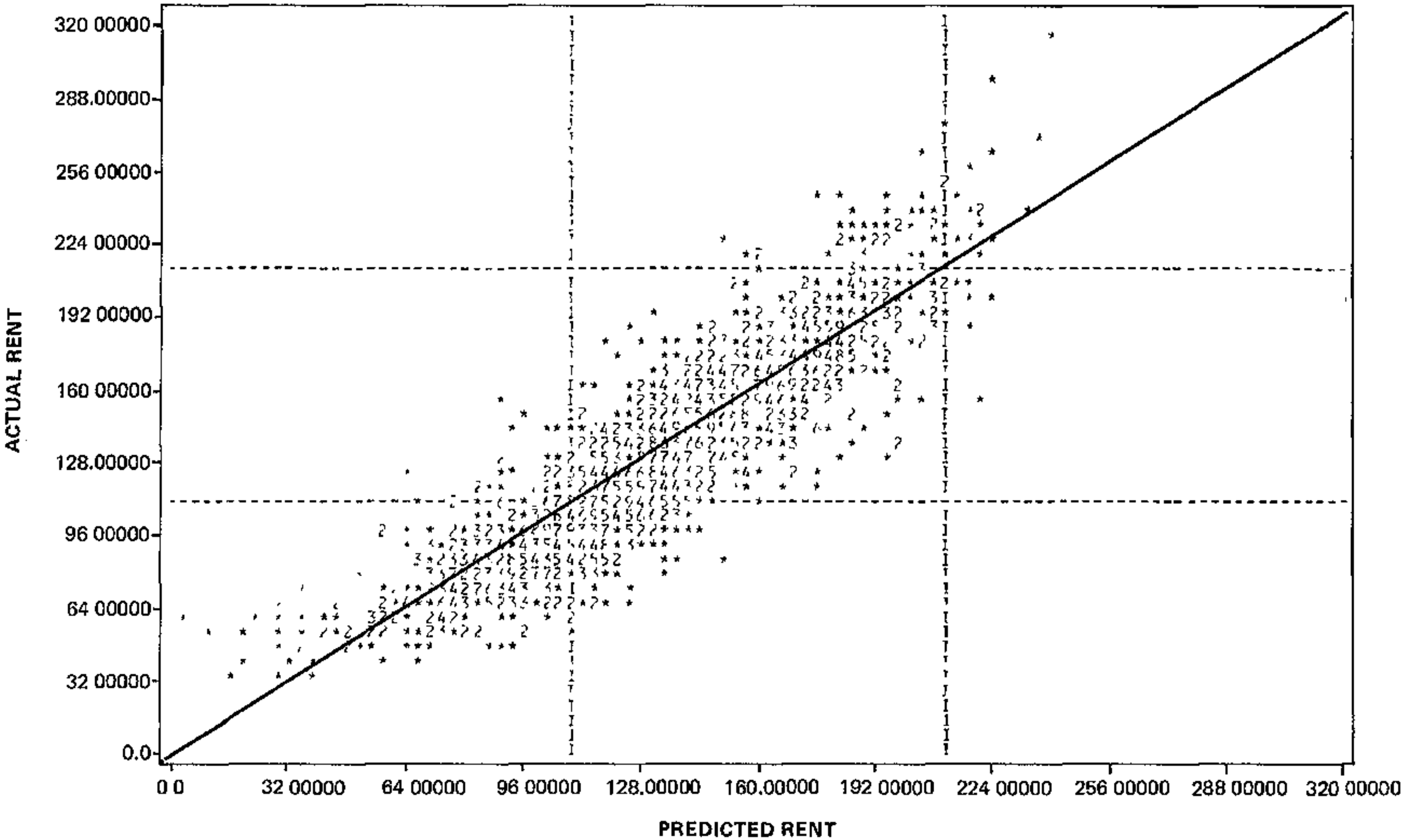
Figure 3-3
ACTUAL AND PREDICTED NATURAL LOG OF RENT: PITTSBURGH



SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

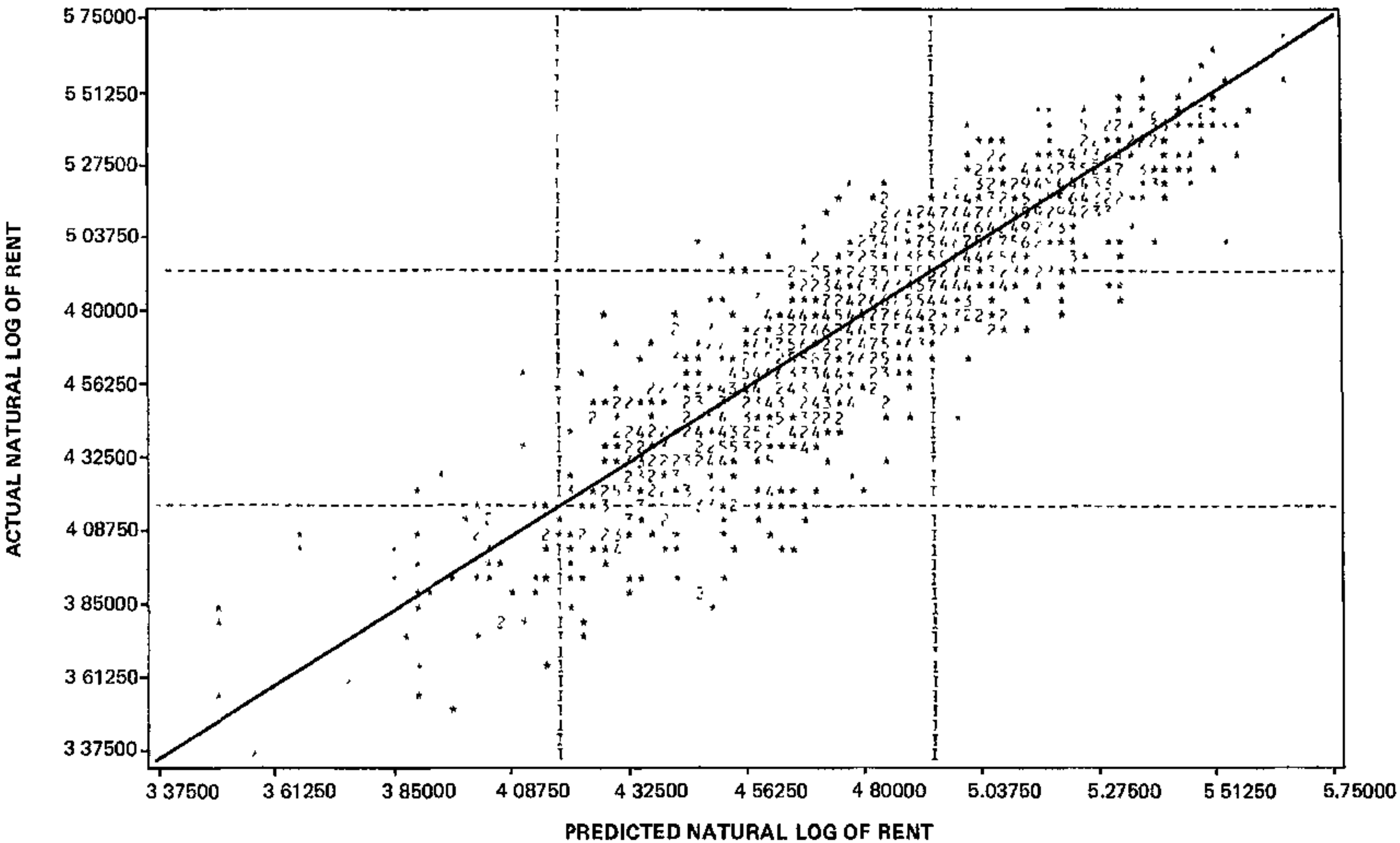
Figure 3-4
ACTUAL AND PREDICTED RENT: PHOENIX



SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

Figure 3-5
ACTUAL AND PREDICTED NATURAL LOG OF RENT: PHOENIX



SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Intervire, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population diagonal line is the locus of points of perfect prediction (that is, equality between actual and predicted values).[1] The observations are tightly distributed around the diagonal for all four equations.

Because the results described here are based on samples from which "outlier" observations have been excluded, one of the more important residuals analyses is the identification of outliers. Following Anscombe (1960) and Anscombe and Tukey (1963), extreme residuals, whether they are felt to represent measurement error or valid observations from a very long-tailed distribution, can be modified or eliminated from formal estimation. When the distribution of residuals was examined, a few cases were felt to be extreme; these outliers were rejected rather than reweighted. Table 3-11 shows the range of residual values for the excluded cases; in the linear equation, for example, the range is -\$106 to \$186 in Pittsburgh, -\$78 to \$201 in Phoenix. (The information for these households was examined extensively and no systematic determinant of poor prediction could be found.) Residuals greater in absolute magnitude than $C\hat{\sigma}$ were excluded, where C is a given constant and $\hat{\sigma}$ is the standard error of the regression (Anscombe and Tukey, 1963).[2] Anscombe and Tukey prefer to view this rejection rule as an insurance policy against a mistaken increase in the error variance. Although the rejection rule was applied to both the linear and semilog equation in the same straightforward manner, the interpretation is not

---

[1]The data points are represented by asterisks (*). When more than one data point falls into a single printing position, the actual number of cases is printed.

[2]C is derived as follows:

$$K = 1.40 + .85N$$

$$C = K\left(1 - \frac{K^2 - 2}{4V}\right)\left(\frac{V}{n}\right)^{1/2},$$

where

$N = 2.326$, the normal deviate for the unit normal cumulative of 0.99

V = degrees of freedom

n = sample size.

Solving C = 3.33: $C\hat{\sigma}$ in Pittsburgh is (3.33)(21.493) = \$71.60 for the linear equation; $C\hat{\sigma}$ in Phoenix is (3.33)(23.934) = \$79.70 for the linear equation. For the semilog equation $C\hat{\sigma}$ equals 0.617 in Pittsburgh, 0.615 in Phoenix. The $\hat{\sigma}$s are the standard errors of the equations prior to exclusion of the outliers.

the same. For the linear case, the rule rejects cases that have residuals with large absolute value. For the semilog case, the residuals are in logarithmic form; thus, cases are rejected in which the percentage of error is large. Cases that were outliers in either equation were excluded from both equations; that is, the union of cases with residual values exceeding $C\hat{\sigma}$ in any equation for a given site were removed from the sample in the final estimation. Fortunately, since the precision of estimation is so similar for the linear and semilog equations for each site, nearly the same cases were defined as outliers in each.[1]

Heteroskedasticity

One of the assumptions of the least squares model is that of homoskedasticity: namely, that the error terms are distributed with constant variance, σ^2. When this condition is not met, the error terms are said to be heteroskedastic. Unless some transformation of the data is done in this case, the least squares estimates will be inefficient.

The exact form that potential heteroskedasticity might take is of course unknown. Several approaches were taken to assess deviations from constant variance. First, the squared residuals were regressed on some of the important independent variables: the number of rooms (natural log); the average of all the surface, structure, and window condition ratings; and the overall evaluation rating. Second, the absolute value of the residuals was regressed on the same independent variables. In addition, as discussed in Section 2.3, the relative degree to which the linear and semilog equations exhibit heteroskedasticity was assessed. In this case, predicted rent was regressed on the squared residuals and on the absolute value of the residuals. These tests are equivalent to assuming, respectively, that the variability of the residuals is proportional to a given independent variable or to the $X\hat{\beta}$ matrix. Finally, scatterplots that correspond to all the regressions were made, in order to provide a visual check on the error terms.

[1] Of the 16 cases excluded in Pittsburgh, 14 were designated extreme values in both equations: that is, both the absolute and proportionate ability to predict were poor. Of the 20 cases excluded in Phoenix, all but 4 cases were excluded by both equations. The difference is largely due to the following: when rents are very low, the percentage of error is sometimes large, whereas the absolute residual is not.

The regression results, shown in Table 3-13, indicate no strong associations between either the squared residuals or the absolute residual and the number of rooms, the average quality rating, or the overall rating. The scatterplots, not presented here, support this conclusion. In most cases the explained variance is close to zero; even when the coefficients are significant, their size is negligible. As discussed in Section 2.2, a very slight heteroskedasticity is present with respect to predicted rent for the semilog equation in Phoenix and for the linear equation in Pittsburgh; however, the scatterplots reveal no regular patterns.

Other Residual Tests

The final set of residuals analyses attempted to uncover systematic biases in predictive ability, according to some geographic pattern or to an association with household characteristics.

First, the mean residuals were computed for Census tracts and for C* neighborhoods; these means were then mapped and coded according to sign and magnitude. No discernible pattern emerged; if access, distance, or neighborhood quality characteristics have been omitted or poorly specified, it is not apparent through this type of test.

Second, the residuals from the linear and semilog equations and the percentage deviation of actual and predicted rent (linear) were regressed on income, education, household size, race, and age of head of household. If important quality attributes have been omitted, the coefficients for income and education would be expected to be large and positive. If racial concentration affects rent, then the coefficient for race of household will be significant if this has not been adequately controlled for. The coefficient might be either negative or positive, depending on whether the race variable proxies omitted neighborhood problems or ghetto price discrimination.[1] No particular hypotheses are maintained for age or size of household.

As shown in Table 3-14, the explanatory power of these equations is very low and the coefficients for most of the variables are small and insignificant. In Pittsburgh, however, the coefficients for income and education are significant and have the expected sign. In Phoenix, income is significant, but education is not. Also, the equation indicates that the residuals for black

[1]This topic is discussed at greater length in Chapters 4 and 5.

Table 3-13

ASSESSMENT OF HETEROSKEDASTICITY

FORM OF RESIDUAL	INDEPENDENT VARIABLE			
	Predicted Rent $\overline{R}^2$	Total Rooms (natural log) $\overline{R}^2$	Average Interior Quality $\overline{R}^2$	Overall Evaluator Rating $\overline{R}^2$
PITTSBURGH				
Linear Equation				
$\|e\|$	.035	.004	.010	.008
e^2	.030	.004	.007	.008
Semilog Equation				
$\|e\|$	.008	.010	.000	.000
e^2	.011	.018	.000	.001
PHOENIX				
Linear Equation				
$\|e\|$	.006	.001	.001	.001
e^2	.010	.004	.000	.002
Semilog Equation				
$\|e\|$	.05	.029	.020	.020
e^2	.045	.000	.01	.02

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 3-14

RESIDUALS ANALYSIS -
REGRESSIONS ON DEMOGRAPHIC VARIABLES

INDEPENDENT VARIABLES AND STATISTICS	DEPENDENT VARIABLES					
	RESIDUALS - LINEAR EQUATION		RESIDUALS - SEMILOG EQUATION		PERCENTAGE DEVIATION OF ACTUAL AND PREDICTED RENT, LINEAR EQUATION	
	Coefficient	t	Coefficient	t	Coefficient	t
			PITTSBURGH			
Income	.00082	2.46	.00001	2.60	-.0000	1.22
Education	.681	3.34	.0059	3.24	-.0004	.24
Household size	.36611	.96	-.0064	1.87	-.0037	1.44
Black household head	-.806	.69	-.006	.54	.0017	.22
Age of head of household	.030	.93	-.0000	.17	.0005	2.47
Constant	-10.70		-.071		.145	
$\bar{R}^2$	.01		.01		.02	
Sample	1,509		1,509		1,509	
			PHOENIX			
Income	.0011	4.18	.00001	3.49	-.00001	4.20
Education	.057	.31	.0010	.69	-.004	3.26
Household size	-.009	.03	-.0011	.42	-.0006	.26
Black household head	-3.49	1.63	-.033	1.96	.0272	1.87
Spanish American household head	-.785	.55	-.002	.20	.012	1.18
Age of head of household	-.054	1.59	.0001	.37	.0005	2.05
Constant	-8.02		-.0448		.1962	
$\bar{R}^2$	.01		.01		.046	
Sample	1,546		1,546		1,546	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Evaluation Form, 1970 Census of Population.

households in Phoenix are on average negative. These results may or may not be meaningful in practice. In Pittsburgh, for example, the linear residuals increase by $0.82 for every $1,000 increase in income. The corresponding figure in Phoenix is $1.10. These amounts seem small. However, the coefficient for black households in Phoenix seems to indicate that variables systematically related to this variable have been omitted.[1]

3.4 ADJUSTMENTS FOR THE COST OF UTILITIES AND FURNISHINGS

The dependent variable used in estimating the hedonic equations is defined as the monthly payment for an unfurnished unit including basic utilities (heat, gas, electricity, water and garbage collection). Since both the rental payment period and whether or not rent includes furnishings and utilities vary among dwelling units, contract rent had to be adjusted to conform to this definition.

Three sorts of adjustments were involved. First, if the payment period was other than monthly, rent paid was adjusted to provide a common (monthly) rental period. Second, if contract rent did not include payments for basic utilities, adjustments were made via site-specific tables for the utilities and services paid for by the household. The utilities involved included heat, gas, electricity, water and garbage collection. The adjustment tables, which are shown in Table 3-15, were derived from interviews with local service and utility companies and public officials and were based on the number of rooms in the dwelling unit.[2] Third, if the unit was furnished, the adjusted rent was reduced by 11.5 percent to deduct for the cost of furnishings. At least one adjustment for some utility or for furnishings occurred for 85 percent of Pittsburgh households and 91 percent of Phoenix households.

An alternative method for taking account of utilities and furnishings would have been to use contract rent as the dependent variable, include the presence or absence of utilities and furnishings in the hedonic regression, and use the estimated coefficients to adjust for utility and furnishings costs.[3]

---

[1] This finding is confirmed by the analysis of discrimination in Chapter 5. In Phoenix, black households appear to pay less than white households for comparison units.

[2] The number of rooms is defined here as number of rooms useable as living space (excluding bathrooms, half-rooms, unfinished basements, or attics). The total number of rooms variable, used in the final equations, includes all rooms.

[3] The results of an estimation using this specification are discussed in Appendix III.

Table 3-15

UTILITY ADJUSTMENT TABLES[a]

NUMBER OF ROOMS IN DWELLING UNIT[b]	1,2	3	4	5	6+
		PITTSBURGH			
Electricity	5	6	7	9	11
Gas	2	2	3	3	4
Heating fuel	10	12	15	18	20
Garbage collection	3	3	3	3	3
Water	3	4	6	7	8
		PHOENIX			
Electricity[c]	11	16	20	24	29
Gas[c]	5	6	7	11	15
Heating fuel[d]	0	0	0	0	0
Garbage collection	3	3	3	3	3
Water	4	4	5	6	9

a. Utility costs are the dollar increments to contract rent per reported utility, by size of dwelling unit. The tables were derived from interviews with local service and utility companies and public officials.

b. The number of rooms is defined as number of rooms useable as living space (excluding bathrooms, half-rooms, unfinished basements or attics).

c. All refrigeration and air-conditioning costs are reflected in the table entries for electricity and gas.

d. In Phoenix, heating fuel costs are assumed to be included in gas and electricity.

The use of adjusted rent instead of contract rent as a dependent variable was based on three considerations. First, the extreme collinearity of the incidence of utility and furnishings adjustments with each other and with unit size makes accurate estimation of a full table of adjustments for all utility and unit size combinations very difficult. Second, since payments were based on adjusted rent, it seemed desirable to maintain the connection between the hedonic index and its associated rent variable and the rent used for payments, if possible. Third, the error introduced by using external adjustments does not in any case seem to pose important problems for either the estimated hedonic index or the analysis of expenditures.

The rest of this section explores the extent and implications of any error in the utility and furnishings adjustments. The results suggest that there was some error. On the other hand, its incidence is low enough and similar enough between Experimental and Control households, and its exact form imprecise enough, to preclude changes in adjusted rent for the analysis of expenditures. Furthermore, its impact on the estimated hedonic value of units is small enough to ignore in estimation of hedonic indices.

An estimate of the error in the adjustments for utilities and furnishings can be obtained using the hedonic model as follows. First, the cost of an unfurnished unit including the cost of basic utilities can be specified as

$$RG = R_c + Z\gamma \tag{5}$$

where

RG = rent including utilities, but excluding furnishings

R_c = contract rent

Z = utilities (furnishings) not included (included) in rent interacted with unit size

γ = true utilities (furnishings) costs.

The analytic rent variable actually used in the regressions (R) is defined as

$$R = R_c + Z\gamma_A = RG + Z(\gamma_A - \gamma) \tag{6}$$

where γ_A are the utility table and furnishings adjustments. Thus analytic rent is true gross rent plus the error, if any, involved in the utility table adjustments.

The hedonic model asserts that true gross rent is a fraction of unit characteristics, so that

$$(7) \qquad RG = X\beta + \varepsilon$$

where X is the matrix of unit characteristics. Any error in utility or furnishings adjustments can be estimated by the equation

$$(8) \qquad R = RG + Z(\gamma_A - \gamma) = X\beta + Z(\gamma_A - \gamma) + \varepsilon \ .$$

Since $\gamma_A > 0$ for utilities, a positive coefficient indicates an over-adjustment $(\gamma_A - \gamma)$. Since $\gamma_A < 0$ for furnishings, a positive coefficient indicates an under-adjustment ($\gamma_A > \gamma$ or $|\gamma_A|<|\gamma|$).

A precise estimate of adjustment error is difficult to obtain for two reasons. First, the incidences of the various individual adjustments are highly collinear. Second, the incidence of these adjustments is strongly related to unit size, as illustrated in Figures 3-6 and 3-7. The greater the number of rooms, the less likely it is that utilities will be included in contract rent (the more likely it is that adjustments from the utility tables were used). For example, fewer than 10 percent of six-room units in either Pittsburgh or Phoenix have payments for electricity included in contract rent, whereas almost 65 percent of one- and two-room units do. This relationship is confirmed by assessing the degree of collinearity among the utility adjustment variables and the other independent variables. The variance inflation factor of the dummy variable for the occurrence of an adjustment was 6.99 in Pittsburgh and 5.67 in Phoenix, indicating that approximately 86 and 82 percent of the respective variances are explained by other variables.[1] For both sites, the high variance inflation factors were primarily due to correlation of the variable for occurrence of an adjustment with the total number of rooms.

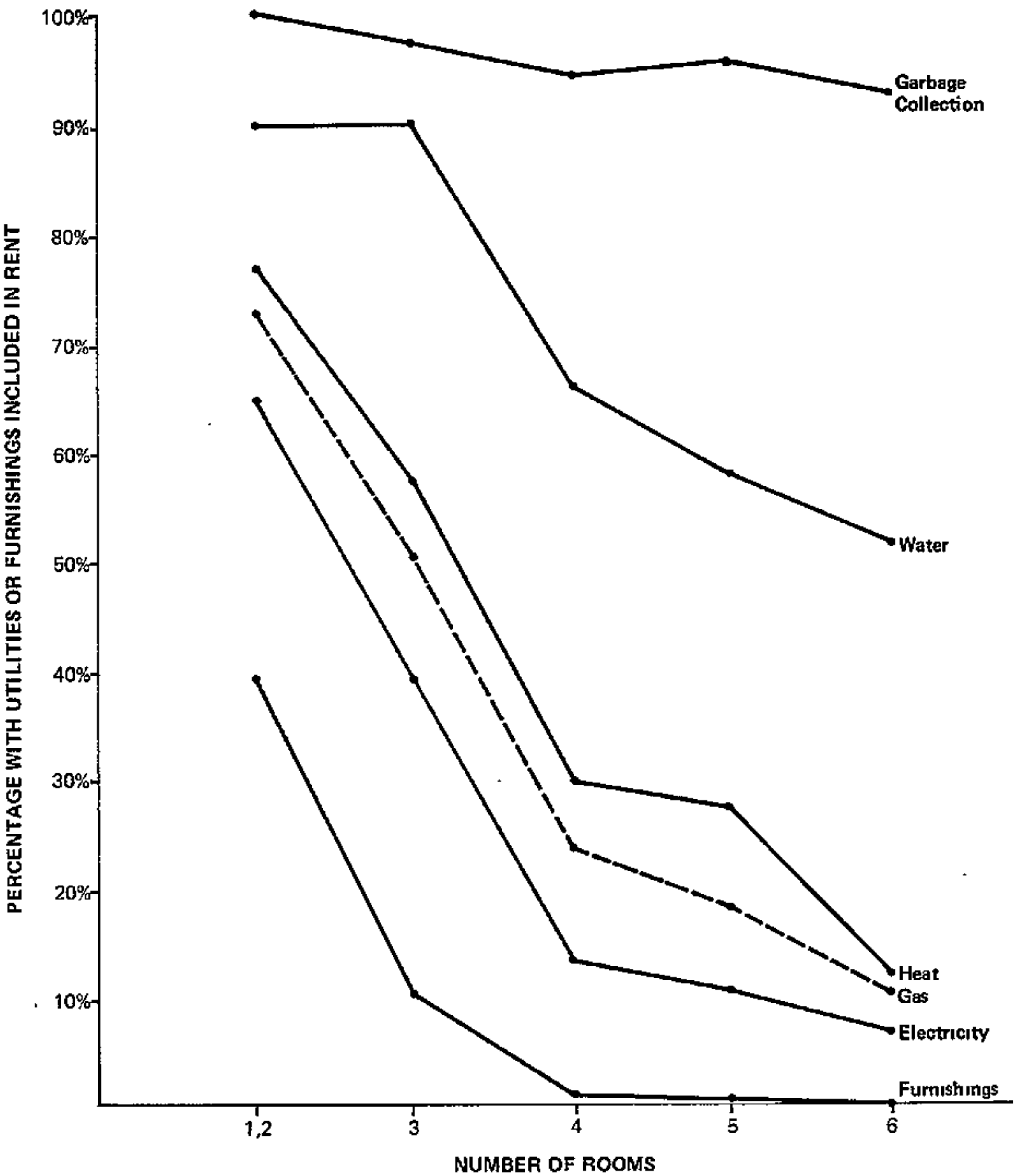
The high degree of collinearity among the utility adjustment variables prevented separate estimations for each utility. As an alternative, many specifications of variables for combinations of utilities and special subsamples were tested. These specifications are summarized in Appendix III, which presents a discussion of the analysis of adjustment errors. Most of the specifications generally produced similar patterns of error estimates based on unit size, although the estimated amount of error differed. Ultimately, two estimates were chosen.

[1] See Section 2.1 for a discussion of collinearity and the definition of the variance inflation factor.

Figure 3-6

TENDENCY OF UTILITIES AND FURNISHINGS TO BE INCLUDED IN CONTRACT RENT, BY DWELLING UNIT SIZE

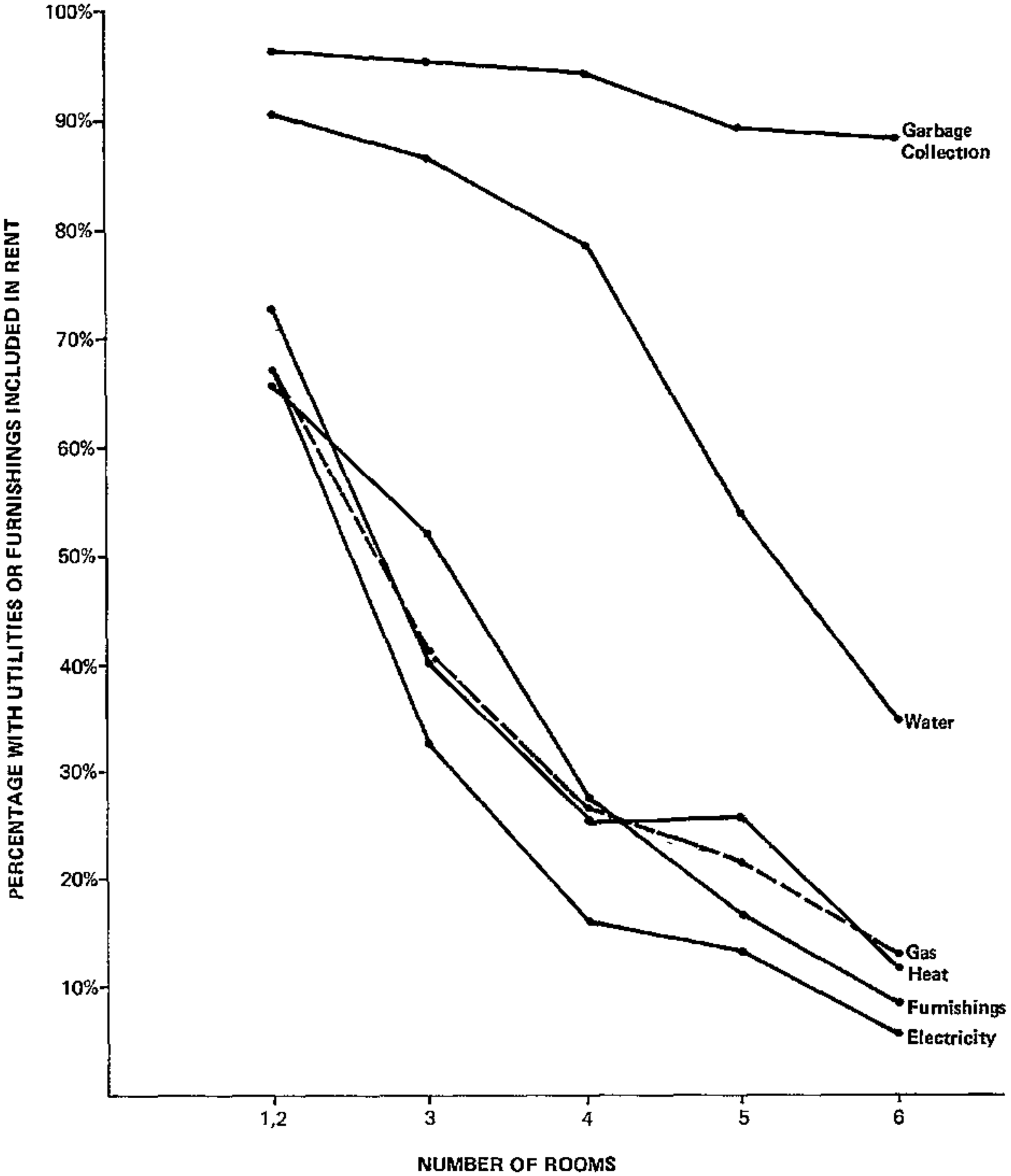
PITTSBURGH



SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in neighborhoods with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES: Initial Household Report Form, Housing Evaluation Form.

Figure 3-7
TENDENCY OF UTILITIES AND FURNISHINGS TO BE INCLUDED IN CONTRACT RENT, BY DWELLING UNIT SIZE
PHOENIX



SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES Initial Household Report Form, Housing Evaluation Form

The first specification provides a general error estimate for the entire population at each site, using a 0-1 dummy variable to indicate the presence of any adjustment for utilities or furnishings and an interaction of the dummy variable with the number of rooms in the unit. Thus the estimating equation for this specification, using the terms from the model developed above, is

$$R_C + Z\gamma_A = X\hat{\beta} + \hat{\delta}_0 Z_0 + \hat{\delta}_1 \ (Z_0 \cdot \text{ number of rooms}) \tag{9}$$

where the dependent variable is analytic rent and $\hat{\delta}_0$, $\hat{\delta}_1$ are estimates of the adjustment error $(\gamma_A - \gamma)$ for the dummy variable Z_0 and the interaction of this variable with unit size. Since over one-third of all Phoenix units are furnished, and since the furnishings adjustment varies depending on the amount of rent paid, a general error was also estimated for Phoenix excluding furnished units. This separate estimate provides an estimate of the utility error independent of the effect of furnishings. The second specification provides estimates of adjustment errors for water, for furnishings, and for the three major utilities together (gas, heat, and electricity). As mentioned above, it was not possible to obtain stable estimates of each of these utilities separately. The sample for this specification is restricted to households that either have adjustments for all three major utilities or no adjustments for any of them.[1] The estimating equation is

$$R_C + Z\gamma_A = X\hat{\beta} + \hat{\delta}_2 Z_2 + \hat{\delta}_3 Z_3 + \hat{\delta}_4 \ (Z_2 \cdot \text{number of rooms}) + \hat{\delta}_5 \ (Z_4 \cdot \text{number of rooms}) \tag{10}$$

where $\hat{\delta}_2$, $\hat{\delta}_3$, $\hat{\delta}_4$, $\hat{\delta}_5$ are estimates of the adjustment errors for the dummy variables Z_2 (gas, heat, and electricity), Z_3 (furnishings), and for the interaction terms of number of rooms with gas, heat, and electricity (Z_2) and with water, Z_4.

The results indicate that adjustment errors did occur and that they were related to unit size. The coefficients of the estimated errors and their standard errors are listed in Table 3-16; many of the coefficients are significant at the 0.01 level. The total estimated errors and their

[1]After testing a variety of variables on a number of subpopulations, it was determined that this restricted sample provided the most reliable error estimate.

Table 3-16

COEFFICIENTS OF ESTIMATED ADJUSTMENT ERRORS

	COEFFICIENT	STANDARD ERROR
PITTSBURGH		
Full Sample		
Any adjustment (0,1)	-4.62	3.80
Any adjustment x number of rooms	2.63**	0.88
Restricted Sample		
Adjustment for gas, heat, and electricity (0,1)[a]	5.36	5.18
Adjustment for gas, heat, and electricity x number of rooms[a]	0.42	1.26
Adjustment for water x number of rooms[a]	0.82**	0.25
Adjustment for furnishings	-4.83	3.46
PHOENIX		
Full Sample		
Any adjustment (0,1)	-33.57**	4.50
Any adjustment x number of rooms	8.56**	1.10
Full Sample Excluding Furnished Units		
Any adjustment (0,1)	-25.37**	5.16
Any adjustment x number of rooms	7.16**	1.25
Restricted Sample		
Adjustment for gas, heat, and electricity (0,1)[a]	-12.02*	5.54
Adjustment for gas, heat, and electricity x number of rooms[a]	3.21*	1.48
Adjustment for water x number of rooms[a]	0.27	0.36
Adjustment for furnishings (0,1)	-8.69**	1.52

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

NOTE: Refer to Appendix III for a discussion of the sample selections and variable specifications.

a. The signs of the coefficients are reversed here, from the equations in Appendix III, to indicate the error for units with adjustments (the equations in Appendix III indicated the value for units with no adjustments). Since the restricted sample contains only households with adjustments for gas, heat, and electricity and those with no adjustment, the populations referenced in the (0,1) dummy variables are opposite and the signs of the coefficients can be reversed.

** Significant at the 0.01 level.

* Significant at the 0.05 level.

relative size with respect to mean analytic rent for each unit size are presented in Table 3-17.

In Pittsburgh, both the general error estimate (full sample) and the total error based on specific utility and furnishing variables (restricted sample) increase with unit size. For the general error estimate, the dollar value and the size of the error as a proportion of mean rent both increase with unit size. For the total error based on specific utility and furnishings variables, the dollar value increases slightly with unit size and the error as a proportion of mean rent remains relatively constant (the increase with unit size for the general error term undoubtedly reflects an increase in the mean number of utilities adjusted for as unit size increases).

In Phoenix, the error estimates are very small for moderate sized units, and increase for both very small and very large units. Thus, the error estimates for four-room units, which contain the greatest number of households in the sample, are very small. Dollar estimates of error were \$0.67 (full sample), \$0.82 (restricted sample), and \$3.27 (full sample exluding furnished units). The largest value for the ratio of estimated error to mean rent was only 0.02 (full sample excluding unfurnished units). The error estimates for three- and five-room units, which contain the next largest number of households, were also relatively small. The largest error estimates occurred for very small (two-room) units or large (six-room) units.

The large negative error for two-room units in Phoenix is partially explained by the furnishings adjustment. Most of the two-room units in Phoenix are furnished and the estimated furnishings error is relatively large and negative, indicating that there was an overadjustment for furnishings. When the general error is estimated on the sample of Phoenix unfurnished units only (see Table 3-17), the estimate is reduced. It still appears, however, that the utility table adjustments in Phoenix underadjusted utility costs in very small units and overadjusted for very large units.

While these figures indicate that adjustment errors were made, they should not be taken too literally, nor do they provide a strong basis for a more correct adjusted contract rent for use in analysis. This proceeds from

Table 3-17

ESTIMATED ERROR IN ADJUSTMENTS, BY DWELLING UNIT SIZE

DWELLING UNIT SIZE (rooms)[a]	2	3	4	5	6
PITTSBURGH					
Adjustment error estimated from general adjustment variable, full sample					
Total estimated error ($)[b]	0.66	3.30	5.94	8 58	11.22
Error as percentage of mean rent	0.91	3 60	5 70	7.00	8.60
Adjustment error estimated from individual utility variables, restricted sample					
Estimated error in gas, heat, and electricity ($)	6.20	6.62	7.04	7.46	7.88
Estimated error in gas, heat, electricity, and water ($)[c]	7.84	9 08	10.32	11.56	12.80
Total error as percentage of mean rent	11.20	10.10	10.10	9 40	9.80
PHOENIX					
Adjustment error estimated from general adjustment variable, full sample					
Total estimated error ($)[b]	-16.45	-7.89	0.67	9 23	17.79
Total error as percentage of mean rent	21.90	7.40	0.48	6.20	10 70
Adjustment error estimated from general adjustment variable, full sample excluding furnished units					
Total estimated error ($)	-11.05	-3.99	3.27	10.43	17.59
Total error as percentage of mean rent	15.10	3.60	2.30	7.00	10.30
Adjustment error estimated from individual utility variables, restricted sample					
Estimated error in gas, heat, and electricity ($)	-5.60	-2.39	0.82	4.03	7.24
Estimated error in gas, heat, electricity, and water ($)[c]	-5.06	-1.58	1.90	5.38	5.62
Total error as percentage of mean rent	6.80	1.50	1 40	3.60	3.40

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Very small (one-room) and very large (seven or more rooms) units are excluded due to small sample sizes.

b. A positive value indicates an overadjustment; a negative value indicates an underadjustment

c. The error in furnishings adjustment is excluded from the sum of errors. Furnishings adjustments occur most frequently for small units and rarely for large units. The estimated coefficient for furnishing error may not be very accurate; thus, the estimate was not added to error estimates for any unit sizes.

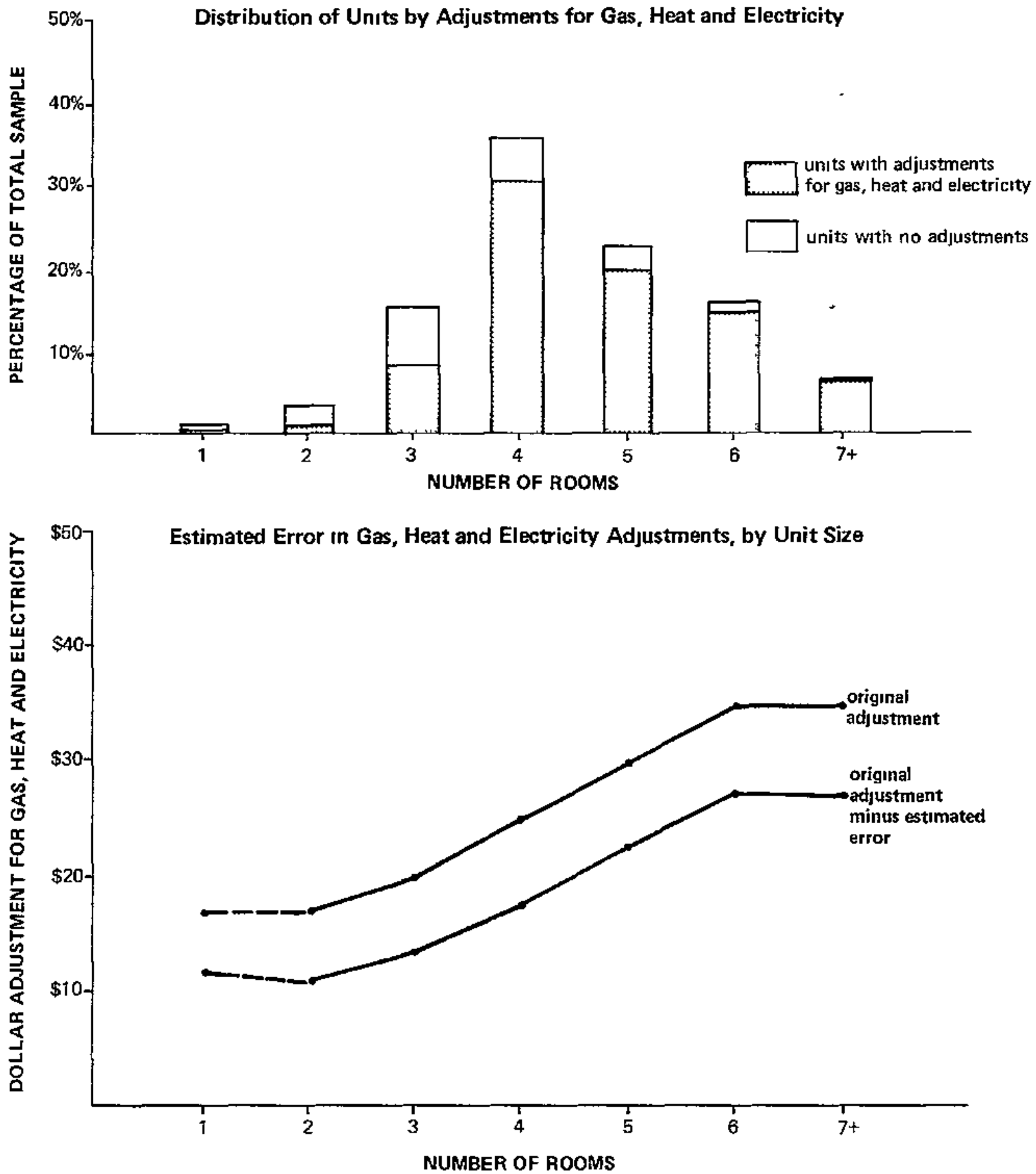
three facts. First, because of the high correlations among the incidence of utilities and furnishings adjustments, separate estimates could not be obtained in either site for gas, heat, and electricity. Thus, there are no estimates appropriate for households that require adjustments on some but not all of these utilities. Second, estimates of the error in the furnishings adjustment in Phoenix are probably not very accurate. Due to the association among unit size and incidence of utility and furnishings adjustment, the estimates fluctuated a great deal when the sample or variable specifications were changed. Third, the utility estimates presented here are based on the presence of adjustments interacted linearly with the number of rooms. In fact, the extremely high and low incidence of adjustments for large and small units and the relatively small number of such units suggest that extrapolation to these unit sizes may be inaccurate. In fact, estimates of adjustment error interacted with unit size dummies are relatively unstable. An F-test indicates that this specification does not have significantly higher explanatory power than the linear interaction for the estimates of gas-heat-electricity adjustment errors using the restricted sample.

Figures 3-8 and 3-9 provide a visual representation of the proportion of units of a given size in the restricted sample, the proportion of each unit size group needing adjustment for gas, heat, and electricity, and the estimated adjustment error. The graphs under the histograms plot the amounts of the utility table adjustments for gas, heat, and electricity and also the "estimated" adjustments--the table amounts less the estimated errors. In Pittsburgh, as discussed above, the overadjustment is not large and is roughly constant across unit sizes. In Phoenix, the impact of adjustment error is greatly attenuated because the unit sizes occurring most frequently in the sample are those with small estimated errors.

The errors in the utility adjustments are not expected to have any serious impact on the analysis. Assuming that the general pattern of the estimated errors is reasonably appropriate, three points should be noted. First, the coefficient for the interaction of adjustment error with number of rooms is relatively small compared to the coefficient for total rooms.[1] Second, if there are errors in the adjustments, they occur equally for Experimental

[1] See Appendix III tables.

Figure 3-8
DISTRIBUTION OF UNITS AND ESTIMATED ADJUSTMENT ERROR BY UNIT SIZE
PITTSBURGH

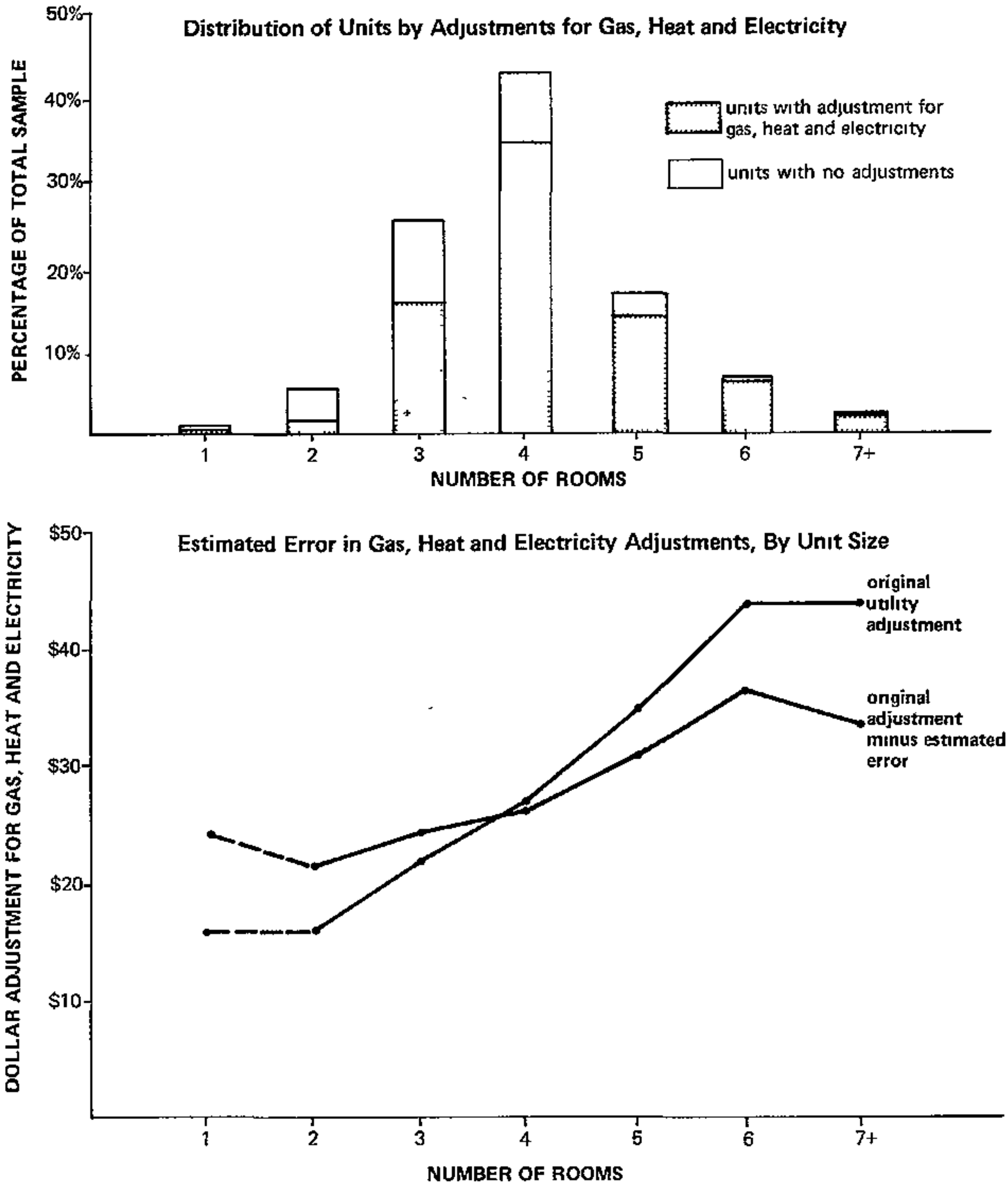


SAMPLE. All enrolled households who have no additional payments for the three major utilities — gas, heat, electricity — or who pay extra for all three utilities, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES: Initial Household Report Form, Housing Evaluation Form

NOTE Dotted line indicates that there were less than 20 one-room dwelling units in the sample

Figure 3-9
DISTRIBUTION OF UNITS AND ESTIMATED ADJUSTMENT ERROR BY UNIT SIZE
PHOENIX



SAMPLE All enrolled households who have no additional payments for the three major utilities — gas, heat, electricity — or who pay extra for all three utilities, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES Initial Household Report Form, Housing Evaluation Form.

NOTE Dotted line indicates that there were less than 20 one-room dwelling units in the sample.

and Control households. Table 3-18 indicates that the percent of Experimental and Control households that have adjustments made to rent is basically the same within each site. Finally, as discussed below, the addition of adjustment variables does not materially affect the estimated hedonic index.

Table 3-18

INCIDENCE OF ADJUSTMENTS, BY TREATMENT TYPE

TYPE OF ADJUSTMENT MADE	PERCENTAGE OF HOUSEHOLDS WITH ADJUSTMENT			
	PITTSBURGH		PHOENIX	
	Experimental	Control	Experimental	Control
Furnishings	5	5	32	34
Gas, heat, and electricity	66	64	64	62
Gas, heat, or electricity	17	19	15	17
No adjustment for gas, heat, or electricity	17	17	21	22
N	1,140	474	1,058	485

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form

The effect of the errors in adjustment in the estimated hedonic index may be assessed by comparing the hedonic equation ($R = X\beta + \varepsilon$) against estimates stratified first by whether or not households had any adjustments and second by whether or not adjustments occurred for all or none of the three major utilities. The results are presented in Tables 3-19 and 3-20. The F-tests, which are significant at the 0.01 level, indicate that the coefficients for

Table 3-19

TESTS FOR SAMPLE DIFFERENCES BETWEEN UNITS ADJUSTED AND NOT ADJUSTED FOR UTILITY OR FURNISHINGS COSTS

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
	PITTSBURGH				
Any adjustment	1,364 (k=32)	.66	.65	467854.03	18.74145
No adjustment	209 (k=32)	.69	.64	70777.40	19.99681
Full sample	1,573 (k=32)	.67	.66	561171.82	19.08300
Tests					
F-test				1.920	
Weighted standard error of subsamples				18.88675	
Comparison of standard errors				.0103	
	PHOENIX				
Any adjustment	1,389 (k=25)	.79	.78	581618.43	20.64961
No adjustment	135 (k=25)	.87	.84	57560.52	22.87526
Full sample	1,524 (k=25)	.79	.79	663291.89	21.03544
Tests					
F-test				2.224	
Weighted standard error of subsamples				20.81685	
Comparison of standard errors				.0103	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 3-20

TESTS FOR SAMPLE DIFFERENCES BETWEEN UNITS ADJUSTED AND NOT ADJUSTED FOR UTILITY COSTS

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		PITTSBURGH			
Adjustments made for gas, heat, and electricity	1,042 (k=33)	.66	.65	325750.04	17.96787
No adjustments made for gas, heat, and electricity	253 (k=33)	.68	.64	89720.57	20.19457
Full sample	1,295 (k=33)	.64	.64	445262.23	18.78357
Tests					
F-test			2.67		
Weighted standard error of subsamples			18.37881		
Comparison of standard errors			.0215		
		PHOENIX			
Adjustments made for gas, heat, and electricity	970 (k=25)	.78	.77	405919.70	20.72546
No adjustments made for gas, heat, and electricity	323 (k=25)	.83	.81	143321.27	21.93043
Full sample	1,293 (k=25)	.79	.78	577505.49	21.34118
Tests					
F-test			2.56		
Weighted standard error of subsamples			21.01218		
Comparison of standard errors			.0154		

SAMPLE: All enrolled households who have no additional payments for gas, heat, and electricity or who pay extra for all three, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

the two populations are different.[1] However, tests for comparison of standard errors indicate that there is little difference in predictive power; the maximum difference in standard errors was only 2 percent.[2]

It is evident, then, that although errors may have occurred in the adjustments, they do not materially affect the estimating ability of the hedonic equation. When errors occur they are most prevalent for small proportions of the population at each site. Furthermore, any errors which exist occur equally for Experimental and Control households. Therefore, the use of analytic rent as the dependent variable does not materially bias the results of the hedonic equations.

---

[1]The critical value of the F statistic being used is $F_{.99}$ (27,1000) 1.738. Most tables do not present the critical F value for samples this large. The value of F cited above was calculated by interpolating for the degrees of freedom in both the numerator and denominator. The values chosen (27,1000) are generally representative of some of the sample sizes in equations below.

[2]The exact procedure used to compare the difference in standard errors of the subsample equations is discussed in Chapter 4.

REFERENCES

Abt Associates Inc., Working Paper on Early Findings, Cambridge, Mass., January 1975.

Alonso, William, Location and Land Use, Cambridge, Mass., Harvard University Press, 1965.

Anscombe, F.J., "Rejection of Outliers," Technometrics, vol. 2, no. 2, May 1960, pp. 123-147.

Anscombe, F.J. and John W. Tukey, "The Examination and Analysis of Residuals," Technometrics, vol. 5, no. 2, May 1963, pp. 141-160.

Draper, N. and H. Smith, Applied Regression Analysis, New York, Wiley and Co., 1966.

Johnston, J., Econometric Methods, New York, McGraw-Hill, 1972.

Kain, John and John Quigley, "Measuring the Value of Housing Quality," Journal of the American Statistical Association, vol. 65, June 1970a, pp. 532-48.

____________________, "Evaluating the Quality of the Residential Environment," Environment and Planning, vol. 2, 1970b, pp. 23-32.

King, Thomas, Property Taxes, Amenities, and Residential Land Values, Cambridge, Mass., Ballinger Publishing Co., 1973.

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

CHAPTER 4

HOUSING MARKETS, HOUSING SUBMARKETS, AND THE USE OF HEDONIC INDICES

This chapter discusses the use of hedonic indices to compare housing services across different cities, across time, and across submarkets within a city. The specific issues raised for Demand Experiment analysis are, first, comparison of the separate Pittsburgh and Phoenix indices with a combined site index; second, comparison of the baseline period equation with an equation estimated after two years; and third, assessment of submarkets based on central city and suburban differences and on racial segmentation.

The analysis of housing services in the Demand Experiment will rely primarily on a market-wide hedonic index applicable to all participants in each city. Thus, major attention has been given to the estimation of separate hedonic regressions in Pittsburgh and Phoenix. These regressions are based on baseline observations and are used to evaluate housing change over the two years of the experiment. Comparison of households in Pittsburgh and Phoenix using the separate site indices may be made in terms of the percentage change in housing services. In some cases, however, a common index for both sites is more desirable.

This chapter assesses the reasonableness of this approach, namely: primary reliance on separate market-wide housing indices for Pittsburgh and Phoenix; the use of baseline period weights to form the index; and the development of a common site index. Whether assessing the index over cities or over time, the central issue is: Do structures of attribute prices differ widely? If market cost structures differ across cities or across time, the hedonic indices associated with each market or time period cannot be directly compared. As discussed in Chapter 1 and Appendix I, hedonic indices cannot reasonably claim to identify a universal index, in terms either of basic consumer preferences or long-run national supply costs. Rather, hedonic indices are market specific; they use the market valuation of units to provide a common ground for indexing housing services within that market.

If relative hedonic weights differ between markets, no clear decision can be made as to which weights should be used.[1]

It should be noted that the real concern is with relative weights. If the weights in one market are simply proportional to the weights in another, then there is no comparison problem--indeed, hedonic indices are often used to correct for this sort of simple inflation. Thus, for example, if housing costs are generally higher in Phoenix, the weights for the regression of rent on attributes will all be inflated proportionally. Using Pittsburgh weights (or combined site weights) corrects for this by adopting a common scale for both sites. Thus a common site index may be used if the hedonic weights differ only by a factor (for the linear form) or by a shift term (for the semilog form).[2]

These issues also arise within the two sites. Recent research indicates that there may be substantial variation in the cost gradients for housing attributes within any given urban area (Strazheim, 1973; Quigley, 1973; Schnare and Struyk, 1974; King, 1973; Kain and Quigley, 1975). If market segmentation exists, indices based on the different submarkets could give different estimates of the change in quality.

Likewise, if the hedonic weights change over time, there is no comparison problem as long as the change in housing costs is proportional: the use of baseline weights simply controls for inflation. If the relative weights

---

[1] In fact, as discussed in Appendix I, the weights of one market may be used to index another. This will tend to overvalue housing in the second market relative to the first.

[2] Note that the market involved is the housing market, rather than, for example, the market for ln(housing). Thus, for the logarithmic form, proportional changes show up as a shift in the constant. Thus if

$$R(\text{Pittsburgh}) = \exp(\beta_0 + X\beta)$$

and

$$R(\text{Pittsburgh}) = \alpha R(\text{Phoenix})$$

then

$$R(\text{Phoenix}) = \exp(\ln\alpha + \beta_0 + X\beta)$$

For further discussion, see Appendix I.

are very unstable, however, then baseline weights provide only an approximate index (see Appendix I).. Indeed, substantial instability would raise questions about the usefulness of the hedonic weights. If the weights change radically within as short a period as two years, it would suggest either that the estimates are unreliable or at least that the weights applicable to any household will depend on the exact time that it moved to its current dwelling unit.

Tests for Market Differences

Two types of tests are used to assess differences in the regression coefficients for different markets or submarkets: F-tests are used to assess the statistical significance of differences, and direct comparison of standard errors is used to assess the operational importance of differences. Loosely speaking, these tests ask whether the subsample regressions perform significantly and substantially better than the full-sample regressions.

First, an F-test is used to test the hypothesis that the overall structure of attribute costs is the same in the two markets or submarkets. If the test is significant, it indicates evidence of differences between markets. Another F-test is made to determine whether the difference in rents in the two markets is only a proportional shift, since application of hedonic indices is not affected in that case.

If evidence of significant market differences is found, it is still important to assess the size of the difference. This is done by comparing the standard errors of the regressions to assess the relative accuracy of prediction using a common equation (as opposed to separate equations) for the various markets or submarkets. Thus, if the F-tests indicate that a significant difference exists, comparison of the standard errors helps assess how important such a difference will be. This is particularly useful when the samples are as large as those in the present study. For large samples, even trivial differences may be statistically significant. If the difference in standard errors is rather small, however, the "error" in prediction made by using the full sample may also be acceptably small.

When appropriate, these tests are also applied at the level of individual coefficient estimates. Again, some judgment should be made about which differences are important. The application of each of these tests is discussed in more detail below.

Statistical tests. In general, the proposition being tested is that the weights in the two markets are the same, or at least that they differ only by a factor (for the linear form) or by a shift term (for the semilog form). This amounts to a test of restrictions on the coefficients for the two markets. Such restrictions may be tested statistically by using the standard F-test on the change in the sum of the squared residuals (SSR) incurred by the restriction--that is,[1]

$$F_{(t,n-r)} = \frac{SSR(R) - SSR(U)}{SSR(U)} \cdot \frac{n-r}{t} \tag{1}$$

where

SSR(R) = the sum of the squared errors estimated under the restriction

SSR(U) = the sum of the squared errors estimated without the restriction

r = the number of parameters in the unrestricted equation

t = the number of restrictions

n = the number of observations.

Thus, for the semilog equation, the restricted specification under the hypothesis that both markets have the same coefficient is

$$\ln R = X\beta + \varepsilon, \tag{2}$$

where X is the vector of independent variables; Equation (2) is estimated for the combined sample from both markets. The unrestricted specification is

$$\ln R = X\beta + \lambda X\delta + \varepsilon, \tag{3}$$

where $\lambda = 1$ for one submarket and zero otherwise. Note that λX only appears for households in one submarket, so that δ represents the difference between the coefficients in the two submarkets. This specification is equivalent to

---

[1] Equation (1) follows from Cochran's theorem and may be derived from, for example, Johnston (1972), pp. 155ff.

stratifying the estimation by submarket. The F-test for the existence of a submarket tests the hypothesis that the vector δ is zero, and is computed as

$$(4) \qquad F_{(k,n-2k)} = \left[\frac{SSR(2)}{SSR(3)} - 1\right]\left[\frac{n-2k}{k}\right],$$

where SSR indicates the sum of squared residuals for the subscripted equation, "n" is the number of observations, and "k" is the number of parameters ("k" equals the rank of the X matrix, that is, the number of independent variables plus the constant term).[1]

Since the use of hedonic indices is not affected when submarket differences are due to a proportional shift in the cost of housing, the semilog of Equation (3) may be compared with one that allows only a simple shift term for submarkets:

$$(5) \qquad \ln R = X\beta + \lambda\delta_0 + \varepsilon,$$

where λ is a dummy variable that indicates one of the two submarkets. Thus, for semilog equations the test statistic,

$$(6) \qquad F_{(k-1,n-2k)} = \left[\frac{SSR(5)}{SSR(3)} - 1\right]\left[\frac{n-2k}{k-1}\right],$$

tests whether stratified estimation by market is superior to using a simple shift term.

For the linear equation, the overall test for submarket is equivalent to that given in Equation (4). Proportional shifts are more difficult, because the restricted regression is nonlinear in its parameters--that is, the specification is:

$$(7) \qquad R = \begin{pmatrix} X_1 \\ aX_2 \end{pmatrix} \beta + \varepsilon,$$

where subscripts refer to observations in the two markets, and "a" is the factor of proportionality. In fact, however, Equation (7) is not estimated; a more relevant test, therefore, is whether the true coefficients differ

---

[1]The value of SSR(3) is often calculated by adding the SSR from the equations estimated separately for the two markets, rather than by directly estimating Equation (3). The two are computationally equivalent.

from the pooled estimates by more than one factor. This may reasonably be tested by the F-test defined by Equation (1). Thus using for the restricted regression:

(8) $$R = \alpha_1 (1-\lambda) \hat{R} + \alpha_2 (\lambda) \hat{R} + \varepsilon$$

where

R = observed rent

$\hat{R}$ = predicted rent using the $\hat{\beta}$ from the pooled equation

λ = a dummy variable taking the value one for one of the submarkets, zero otherwise.

and using for the unrestricted linear regression,

(9) $$R = X\hat{\beta} + \lambda X\delta + \varepsilon,$$

a reasonable test statistic for the linear equation might be:[1,2]

(10) $$F_{(k-1,n-2k)} = \left[\frac{SSR(8)}{SSR(9)} - 1\right]\left[\frac{n-2k}{k-1}\right].$$

In the case of differences over time, the test is made against the baseline coefficients, so that the $\hat{R}$ in Equation (8) is taken from these estimates alone, rather than from pooled estimates for both time periods.[3]

---

[1] Note that the F-test of Equation (10) may not apply exactly, since it is not immediately clear that the estimated standard error from Equation (9) will be distributed independently of the coefficient estimates in Equation (8).

[2] The degrees of freedom in Equation (10) reflect the fact that "k" degrees of freedom were used up in estimating $\hat{R}$ for Equation (8), and also that although two parameters are estimated in Equation (8), one reflects an arbitrary normalization of $\hat{R}$.

[3] The reason for using the baseline weights alone is because the additional sample which could be used for estimation at the later time period is small. The hedonic index is used to test for changes in the rent/quality relationship induced by the experiment. Thus, observations on the experimental households cannot be used to estimate the basic hedonic coefficients except at baseline (before households were informed of the experiment). The sample of observations at the end of two years is therefore restricted to Control households.

Assessing the Importance of Differences

For very large samples, even trivial differences in coefficients will be statistically significant. Thus it is important to assess not only whether the observations deviate from the restrictions in a statistically significant way, but also whether those differences are operationally worth considering. This of course depends on the values of the independent variables. If, for example, the difference coefficients (δ) in Equation (3) are not zero, then, for large enough values of X, the term $X\delta$ (the difference between the two markets) will be large. The problem in assessing operational importance is therefore to arrive at some reasonable range of values for the dependent variable. One convenient summary measure is the difference in the standard error of the regression estimated with and without the restriction --that is,

$$\emptyset = \frac{SE(R) - SE(U)}{SE(R)} \tag{11}$$

where

SE(R) = the standard error of the restricted regression

SE(U) = the standard error of the unrestricted regression.

This measure, proposed for hedonic indices by Ohta and Griliches (see Ohta and Griliches, 1972; Schnare and Struyk, 1974) in effect weights the differences in coefficients according to the incidence of various values of the independent variable in the sample. This is not intended as an exact test. It does, however, offer a convenient summary statistic on the overall impact of coefficient differences.[1]

---

[1]There is of course a close connection between the change in standard errors and the F-test of Equation (1). As normally used, the change in standard error in $\emptyset$ is expressed in terms of the reduction in the standard error when the restriction is relaxed--that is,

$$\emptyset = \frac{SE(R) - SE(U)}{SE(R)} , \tag{a}$$

where R and U refer to restricted and unrestricted estimates, respectively, and the standard error for a regression is defined by

$$SE = \left(\frac{SSE}{n-r}\right)^{1/2} , \tag{b}$$

(footnote continued on next page)

Another way to assess the magnitude of differences is to examine the individual differences in coefficients. Thus, in a form such as Equation (3), the differences in coefficients between the two markets are directly estimated by estimates of δ. These may be tested individually, and their operational importance can be assessed in terms of the size of the coefficient and the variation in the relevant independent variable.

Summary of Findings

Section 4.1 discusses the combined site hedonic index and compares it with the independently derived equations for Pittsburgh and Phoenix. Tests for market segmentation due to central city/suburban differences and to racial and ethnic differences are addressed in Section 4.2. Finally, Section 4.3 examines the hedonic regressions for the two-year time period. The following conclusions emerge:

> Market Differences. The explanatory power of the common variable set is good and the index appears reasonable. The structure of attribute prices does differ across Pittsburgh and Phoenix, and the coefficients for a few

---

(footnote continued from previous page)

where SSE is the sum of squared errors, "n" is the number of observations, and "r" is the number of free parameters in the regression.

The F-test usually uses the SSE from the unrestricted regression in the denominator, though this is somewhat arbitrary. In any case, recognizing that in terms of Equation (1) the number of free parameters is "k" for the unrestricted and "k-t" for the restricted, a little algebra will show that

$$\text{(c)} \qquad \frac{1}{(1-\emptyset)^2} = \left(1 + F\frac{t}{n-k}\right)\left(1 + \frac{t}{n-k}\right).$$

Thus the standard error measure is a monotonic transform of the F-statistic for the regression. Indeed, when $\hat{F}$ is zero, then

$$\text{(d)} \qquad \emptyset(F=0) = 1 - \sqrt{\frac{n-k}{n-k+t}}.$$

This difference in the standard error is solely due to the difference in degrees of freedom. Thus the F-test may be regarded as a test of whether or not $\emptyset$, corrected for degrees of freedom, is zero.

In fact, $\emptyset$ was often used before the F-distribution was tabulated. A corrected-for-bias version may be found in Kelly (1935).

of the important attributes are quite different. Keeping this in mind, the common variable index can be viewed as an approximate, but convenient, way to make direct comparisons of housing change when a pooled sample is appropriate in direct cross-site comparisons.

Central City and Suburban Submarkets. No evidence of market segmentation exists for the central city and suburban submarkets in either Pittsburgh or Phoenix. This may be because the sample consists only of lower-income households, whose neighborhoods may be reasonably homogeneous. In any event, attribute prices in these submarkets are not significantly different, so a single housing index is applicable.

Minority and White Submarkets. Differences between racial or ethnic groups were assessed in two ways: differences between neighborhoods (Census tracts) that are primarily white or primarily minority, and differences between households where the head is white or minority. In both cases, there is only small evidence of submarket differences. Thus, in Phoenix, some segmentation of Spanish American and white neighborhoods (and households) is indicated, but the differences in predictive power between the submarket and full-sample equations are small, ranging from only 0.5 percent to 3 percent.

In Pittsburgh, the F-statistics indicate no difference between equations for ghetto and for white neighborhoods, but significant differences between equations for black and for white households. However, since these submarket equations differ in predictive power from the full-sample equation by less than 1 percent, it is hard to see any practical effect.

Based on these submarket tests, the use of a market-wide equation in each site seems appropriate.

Comparison Over Time. Comparison of the hedonic regressions for the baseline and two-year periods (only Control households were used in estimation of these equations) indicates, as expected, relatively little instability in coefficient prices. Thus, the baseline weights are as efficient as the estimated two-year weights in predicting rent at two years.

4.1 THE COMBINED SITE HEDONIC EQUATIONS

Hedonic indices may be expected to provide only an approximate method for comparing housing services across cities in which attribute costs differ significantly. The analysis in this section concludes that the Pittsburgh and Phoenix cost structures do in fact differ and that these differences are not simply due to proportional shifts. Thus, housing outcomes will generally be described using separately estimated hedonic equations for Pittsburgh and Phoenix. In this context, the percentage change in housing services can be used to compare the change in housing outcomes across cities.

However, when a direct comparison of the level of housing services across cities is desirable, a common index is needed. Thus, combined site equations, both linear and semilog, have been specified. The equations use only those variables that can be defined and are significant in both cities. The specification follows the approach presented in Chapter 2 regarding coefficient test level, collinearity, and so forth.

Two types of questions are raised in comparisons between the pooled estimation and the independently specified regressions. First, what happens to the explanatory power for each city when the common variable list is used instead of the separately derived lists? And, how well does the common variable list describe dwelling unit and neighborhood characteristics? Second, to what extent do the coefficients for dwelling unit and neighborhood variables in Pittsburgh and Phoenix differ?

The overall conclusion is that while separately derived indices can provide somewhat better predictions, the results of the common index are reasonable. Explanatory power suffers some in Pittsburgh, but hardly at all in Phoenix (see Table 4-1). Many of the variables that describe the dwelling unit are the same as or similar to those in the separate equations. Neighborhood quality, however, is not very adequately described in the combined equation, primarily because principal components analysis was not feasible using the pooled sample. A more serious problem is that the Pittsburgh and Phoenix coefficients for a few of the important attributes, such as space, are significantly different. Thus, the coefficients of the common equation must be regarded simply as averages of different underlying cost structures; they cannot support the more consistent "implicit price" interpretation accorded to the coefficients of the separate city equations.[1]

---

[1] Even if submarkets exist within a city, the differences are not likely to be as major as between cities of different climates, regions, and periods of development. A within-city submarket must be "maintained" by some sort of barrier, such as discrimination or extremely lumpy attribute combinations. Certainly, distance helps to maintain the different relative cost structures in Pittsburgh and Phoenix. Also, very few of the attributes (footnote continued on next page)

Table 4-1

COMPARISON OF EXPLANATORY POWER:
SEPARATE SITE AND COMMON SITE EQUATIONS
(Adjusted $\overline{R}^2$)

EQUATION	COMMON SITE VARIABLE LIST			SEPARATE SITE VARIABLE LIST	
	Pooled Sample	Pittsburgh	Phoenix	Pittsburgh	Phoenix
Linear	.71	.57	.77	.66	.79
Semilog	.72	.61	.78	.66	.80

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Market Comparison, Pittsburgh and Phoenix. As discussed above, differences between the two sites were tested in terms of overall significance, significance of nonproportional differences, and operational importance in terms of the percentage change in the standard error. The results are shown in Table 4-2, which presents summary statistics for the following equations: the semilog and linear pooled site equations; the separate Pittsburgh and Phoenix estimations, using the common variable list; a pooled equation, with a dummy variable for Phoenix; and a pooled equation that allows for full interaction between the site (Phoenix in this case) and the attributes.[1]

(footnote continued from previous page)
are portable (like dishwashers or air-conditioners) or are likely to be completely independent of other attributes. Thus, there is little reason to expect the estimated coefficients to be the same in Pittsburgh and Phoenix.

[1]The actual equations may be found in this subsection or in Appendix II. The equations listed in this subsection are the pooled sample semilog equation, the pooled sample linear equation, and the linear equation with full interaction terms between attributes and site. The other equations are presented in Appendix II.

As described previously, an F-test is first performed to test for overall differences between the markets. If a significant difference is shown, another F-test assesses whether the difference is attributable to a proportional shift or to relative price differences. As seen in Table 4-2, both F-tests are significant. The third test presented in Table 4-2, the comparison of standard errors, indicates that a 2 to 3 percent loss of predictive power occurs when using pooled rather than stratified market samples.

The magnitude of the Pittsburgh/Phoenix price difference can be assessed in two ways using the available data. First, a simple specification is to include a shift term for site in the equation estimated using the pooled sample. This approach suggests that housing costs about \$10 (linear equation) or 7.3 percent (semilog equation) more in Phoenix than in Pittsburgh.[1] However, since <u>relative</u> attribute prices differ, it is more appropriate to derive the price indices from the stratified estimations.

This derivation enables estimation of the price of the average Pittsburgh housing bundle if purchased in Phoenix and, correspondingly, the price of the average Phoenix bundle if purchased in Pittsburgh. (Refer to Table 4-3 for these computations.) For example, the average Pittsburgh unit costs \$110.19 (this is the sample mean rent), but would cost \$121.17 if rented in Phoenix.

When the estimates are related to the actual prices of the particular housing bundles, two indices can be formed:[2]

$$(12) \qquad \frac{\Sigma X_{i1}\hat{\beta}_{i2}}{\Sigma X_{i1}\hat{\beta}_{i1}}$$

or

$$(13) \qquad \frac{\Sigma X_{i2}\hat{\beta}_{i2}}{\Sigma X_{i2}\hat{\beta}_{i1}}$$

[1]The equations are listed in Appendix II.

[2]Economists will recognize these as the usual Laspeyres and Paasche indices. In general, these indices tend to over and underestimate a true index, respectively. However, the true index need not be the same in each case.

Table 4-2

COMMON VARIABLE LIST:
PITTSBURGH AND PHOENIX

EQUATION	NUMBER OF OBSERVATIONS	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION		
Pittsburgh common variable list	1617 (k=24)	.61	55.89628	.18732
Phoenix common variable list	1607 (k=24	.78	49.924	.17759
Full sample (pooled sites)	3245 (k=24)	.72	112.621	.18699
Full sample with shift term	3245 (k=25)	.72	111.288	.18591
Full sample with complete interaction	3245 (k=48)	.73	107.317	.18322

Submarket Tests:

F-test for submarket	6.5805 (24,3197)
F-test for proportional shift	5.56 (23,3197)
Comparison of standard errors	0.0235

EQUATION	NUMBER OF OBSERVATIONS	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		LINEAR EQUATION		
Pittsburgh common variable list	1618 (k=19)	.57	722609.105	21.258
Phoenix common variable list	1607 (k=19)	.77	793266.48	22.35035
Full sample (pooled sites)	3229 (k=19)	.71	1621485.00	22.475
Full sample with shift term	3229 (k=20)	.71	1593089.05	22.281
Full sample with complete interaction	3229 (k=38)	.72	1525272.71	21.863
Full sample with proportional shift	3229 (k=38)	.71	1601979.0	22.28

Submarket Tests:

F-test for submarket	10.598 (19,3191)
F-test for proportional shift	4.432 (18,3191)
Comparison of standard errors	0.0296

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households. The sample used for estimation of the separate Pittsburgh and Phoenix equations is slightly smaller than that used for the pooled sample equations. This does not affect the results since the tests are performed using the pooled sample.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-3

COMPARISON OF PRICES FOR MEAN
HOUSING BUNDLES: PITTSBURGH AND PHOENIX

COMMON SITE EQUATION: LINEAR

HOUSING BUNDLE	PITTSBURGH COEFFICIENTS	PHOENIX COEFFICIENTS	PRICE INDEX
Pittsburgh average housing bundle	\$110.19	(\$121.17)[a]	$\frac{\Sigma X_{i1}\hat{\beta}_{i2}}{\Sigma X_{i1}\hat{\beta}_{i1}} = 1.10$
Phoenix average housing bundle	(\$124.05)[a]	\$132.32	$\frac{\Sigma X_{i2}\hat{\beta}_{i2}}{\Sigma X_{i2}\hat{\beta}_{i1}} = 1.07$

COMMON SITE EQUATION: SEMILOG

HOUSING BUNDLE	PITTSBURGH COEFFICIENTS	PHOENIX COEFFICIENTS	PRICE INDEX
Pittsburgh average housing bundle	4.6579 \$105.41	[(4.7354) \$113.91][a]	$\frac{\Sigma X_{i1}\hat{\beta}_{i2}}{\Sigma X_{i1}\hat{\beta}_{i1}} = 1.08$
Phoenix average housing bundle	[(4.7682) \$117.71][a]	4.8180 \$123.72	$\frac{\Sigma X_{i2}\hat{\beta}_{i2}}{\Sigma X_{i2}\hat{\beta}_{i1}} = 1.05$

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Estimated using the mean attribute levels of one site and the coefficients and intercept in the other site. The indices for the semilog equation are formed from the exponential rents.

where X_{ij} = the mean housing attributes values in site

$\hat{\beta}_{ij}$ = the coefficients estimated for site "j."

Based on the linear equation, average housing in Phoenix costs about 7 to 10 percent more than in Pittsburgh; for the semilog equations, the range is 5 to 8 percent.

These estimated differences are of course based on a sample of low-income households and may not be relevant to other segments of the market. Most important, the percentage differences in price will vary as the housing bundles change. For example, the price of space is much higher in Phoenix than in Pittsburgh;[1] thus, for larger units, the cost differences will increase.

Variables Used in the Common Site Equation. As already mentioned, the variables used for pooled site regressions were somewhat different from those used in the individual site estimates presented in Chapter 3. Table 4-4 shows the means and standard deviations of the variables retained in the common equation; this table may be contrasted with the variable lists for the separate Pittsburgh and Phoenix equations, described in Chapter 3 (Table 3-1). Tables 4-5 and 4-6 present the pooled sample equations.

The variables that describe tenure characteristics in the common site equation are generally similar to those in the separate equations. An exception is the "length of residence" variable, which is an exponential function in the separate equations, with different parameters in Pittsburgh and Phoenix. A simpler form of the residence variable--the natural log of the number of months of residence--provides very good fit and was used in the combined equation.

Many of the variables that represent dwelling unit quality and size in the combined equation are similar to those in the individual equations. Other variables use combined attributes in order to reflect some unique site differences. Thus, the "parking facilities" variable represents a garage or carport in Phoenix, and a garage or offstreet parking in Pittsburgh.

---

[1] Refer to the discussion in the following subsection and to Table 4-7 for assessment of coefficient differences.

Table 4-4

MEANS AND STANDARD DEVIATIONS· COMMON SITE VARIABLES

VARIABLE DESCRIPTION		PITTSBURGH MEANS[a]	PITTSBURGH STANDARD DEVIATION	PHOENIX MEANS[a]	PHOENIX STANDARD DEVIATION	SITES POOLED MEANS[a]	SITES POOLED STANDARD DEVIATION
Tenure Charac-teristics		3.472	1.112	2.743	1.022	3.109	1.128
	Related to landlord (0,1)	071	.257	.057	.232	064	246
	Number of landlord contacts for maintenance	1.329	1.390	1.263	1.380	1.295	1.385
	Number of persons per room	.699	335	.840	.468	.769	.412
Dwelling Unit Features	Area per room (natural log)	4.850	.184	4.687	.202	4.768	.209
	Total number of rooms (includes kitchen & bath) (natural log)	1.656	.308	1.589	.258	1.622	.287
	Building age (years)	50 024	14.004	24.462	15.122	37.285	19.386
	Dishwasher and/or disposal provided (0,1)	.053	.224	.159	.366	.106	.307
	Adequate light and ventilation (0,1)	.415	493	.391	488	.403	.491
	Average surface and structural quality (4 point scale)	2.168	.369	2.285	.633	2.226	.522
	Parking facilities provided (0,1)	.148	.355	.312	.463	.230	421
	Large multifamily structure (0,1)	.148	.356	.146	.354	.147	.355
	Recent interior painting or papering (0,1)	.100	.300	204	403	.152	.359
	Many high quality features (0,1)	040	.195	.125	.331	082	.274
	Presence of private yard (0,1)	.364	.481	.518	.500	.441	.497
	*Temperature control· central heat or air-conditioning (0,1)	.576	.494	.344	.475	.461	.499
	*Inferior or no heat (0,1)	.218	.413	.315	465	.265	.442
	*Presence of adequate ceiling height (0,1)	.910	.286	.905	.293	.907	.290
	*Adequate plumbing present and working (0,1)	.830	.376	.838	.369	.835	.371
	*Stove and refrigerator provided (0,1)	.113	.316	641	.480	.374	484
Neighborhood Features	Median income of census tract (dollars)	8502 807	1623.467	8072.137	2148.115	8288.452	1915.310
	Quality of block face landscaping (4 point scale)	1.370	.934	1.699	.823	1.533	.896
	Quality of adult recreation facilities	1.417	.296	1.597	.274	1.506	.299
	Distance from Central Business District (miles)	5.472	3.722	5 376	4.299	5 428	4 020
Rent	Analytic rent	110.186	32.637	132.324	46.411	121.204	41.596
	Natural logarithm of analytic rent	4.658	.302	4 818	.381	4 738	.353

SAMPLE· All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Mean values refer to the sample sizes for the semilog equation N = 1,617 for Pittsburgh and N = 1,607 for Phoenix, for the separate site equations, and 3,245 for the pooled semilog equation Sample sizes for the linear equation: N = 1,618 for Pittsburgh and N = 1,607 for Phoenix for the separate equations, and 3,229 for the pooled linear equation.

* Variables with asterisk are in the semilog equation only.

Table 4-5

POOLED SAMPLE LINEAR EQUATION: COMMON SITE VARIABLES

R^2 = 0 710 $\bar{R}^2$ = 0.708 F = 435.947 N = 3,229

VARIABLE DESCRIPTION		COEFFICIENTS	t-STATISTIC
Tenure Charac-teristics	Length of residence (natural log)	-6.360	16.384
	Related to landlord (0,1)	-13.299	8.106
	Number of landlord contacts for maintenance	1.445	4.932
	Number of persons per room	7.566	7.159
Dwelling Unit Features	Area per room (natural log)	28.097	13.529
	Total number of rooms (includes kitchen & bath) (natural log)	62.293	41 049
	Building age (years)	-0 471	15.382
	Dishwasher and/or disposal provided (0,1)	11.230	7.453
	Adequate light and ventilation (0,1)	4.758	5.395
	Average surface and structural quality (4 point scale)	12.946	11.659
	Parking facilities provided (0,1)	8.521	8.325
	Large multifamily structure (0,1)	9.012	7.152
	Recent interior painting and papering (0,1)	6.596	5.754
	Many high quality features (0,1)	8.937	5.339
	Presence of private yard (0,1)	3.035	3.421
Neighborhood Features	Median income of census tract (dollars)	0.002	8.924
	Quality of block face landscaping (4 point scale)	3.622	7.582
	Quality of adult recreation facilities	18.900	12.844
CONSTANT		-174.057	

SAMPLE All enrolled households, excluding those that moved between Baseline Interview and enrollment those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-6

POOLED SAMPLE SEMILOG EQUATION COMMON SITE VARIABLES

R^2 = 0 721 $\bar{R}^2$ = 0 719 F = 361 592 N = 3,245

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Charac-teristics	Length of residence (natural log)	- 053	16.275
	Related to landlord (0,1)	- 109	7.988
	Number of landlord contacts for maintenance	.014	5 915
	Number of persons per room	.076	8.572
Dwelling Unit Features	Area per room (natural log)	.247	13.816
	Total number of rooms (includes kitchen & bath) (natural log)	.613	46.672
	Building age (years)	-.004	13.632
	Dishwasher and/or disposal provided (0,1)	042	3.302
	Adequate light and ventilation (0,1)	.022	2.996
	Average surface and structural quality (4 point scale)	.089	8.733
	Parking facilities provided (0,1)	.058	6 790
	Large multifamily structure (0,1)	.053	5.094
	Recent interior painting or papering (0,1)	.042	4.449
	Many high quality features (0,1)	.038	2.682
	Presence of private yard (0,1)	N/A	N/A
	Temperature control: central heat or air-conditioning (0,1)	023	2.756
	Inferior or no heat (0,1)	-.068	6.864
	Presence of adequate ceiling height (0,1)	.047	4.000
	Adequate plumbing present and working (0,1)	.018	1.803
	Stove and refrigerator provided (0,1)	.060	6.705
Neighborhood Features	Median income of census tract (dollars)	0	7.410
	Quality of block face landscaping (4 point scale)	.026	6 629
	Quality of adult recreation facilities	.134	10.662
	Distance from Central Business District (miles)	- 003	3 229
CONSTANT		2.116	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Similarly, the "adequate temperature control" variable combines central heat and various air-conditioning variables in a slightly different manner in each site.

Interior surface and structural quality is represented in the combined equation by the average of all the surface structure ratings, the same variable now used in the separate Phoenix equation. In the separate Pittsburgh equation, however, interior quality is represented by derived factor scores. Since the principal components solutions differed widely between the two sites, no attempt was made to perform principal components analysis on the surface and structure ratings in the pooled sample.

The major problem in definition of variables is the very limited specification of neighborhood quality in the common site equation. In the individual Pittsburgh and Phoenix equations, neighborhoods are described primarily by two sets of derived factor scores; one group represents Census tract descriptors, the other represents aggregations of participant opinions of neighborhood services. (Refer to Chapter 3 and Appendix VI for a description of these variables.) Again, since the principal components solutions were quite different in Pittsburgh and Phoenix, a pooled sample factor analysis was not done. Instead, individual Census tract variables and averages of participant ratings were tested in the common equation. Since these sets of variables are each extremely collinear, only a very limited number of neighborhood variables might be expected to be retained in the equation; in fact, only one variable from each set is retained. Census tract median income is used as a proxy for Census tract characteristics. The quality of adult recreational facilities, aggregated across groups of tracts, represents these larger neighborhoods. Finally, accessibility is represented by distance from the Central Business District; this variable is retained only in the semilog equation, however.

Test of Coefficient Differences. Differences between the two sites may also be examined in terms of the individual coefficients. Since Pittsburgh and Phoenix differ along important dimensions, such as mobility rate and the age and size of the housing stock, differences in the estimates of some individual coefficients might be expected. In order to test for significant coefficient differences, the following equation was estimated:

(14) $$R = X\beta + \lambda X\delta + \varepsilon,$$

where λ is a dummy variable that represents Phoenix.[1] Thus δ represents the difference between the Pittsburgh and Phoenix coefficients. These differences (and the corresponding t-tests) are displayed as the interaction variables in Table 4-7.

A substantial number of variables have somewhat different coefficients. In only a few cases, however, are the differences striking and not all of these affect the housing index. For example, based on the length of residence variable in the common (linear) equation, the tenure discount in Phoenix exceeds the discount in Pittsburgh by about $8.75 after two years of residence. This is not surprising, given the different mobility patterns and vacancy rates in these cities. The average discount for being related to the landlord is also greater in Phoenix. Since these variables are not included in the housing index, and since they are reasonably uncorrelated with the quality variables, differences in the tenure variables are not of major concern to the housing index itself.

A more serious problem, however, is the difference between the cost of dwelling unit variables, particularly space, in the two cities. Based on the linear equation in Table 4-7, a three-room apartment costs about $26 more in Phoenix than in Pittsburgh; a four-room unit costs $33 more. Large differences also occur in the cost of square feet per room.

Interior quality (as measured by the average of surface and structure ratings) also costs more in Phoenix than in Pittsburgh. The markup for "many good features" is less in Phoenix than in Pittsburgh, however. The premium paid for parking facilities or a large multifamily unit is less in Phoenix than in Pittsburgh.[2] Interestingly, the cost of a dishwasher or disposal--a portable attribute--is the same in both cities. This would be expected for these attributes are directly produced by the market and (except for transportation costs) prices would be expected to be the same in the two sites.

---

[1] The semilog equation that corresponds to Equation (14) in the text is listed in Appendix II.

[2] Recall that the definition of parking facilities differs in the two sites.

Table 4-7

POOLED SAMPLE LINEAR INTERACTION EQUATION. COMMON SITE VARIABLES

R^2 = 0 727 $\bar{R}^2$ = 0 724 F = 229.566 N = 3,229

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Length of residence (natural log)	-4.738	9 303
	Related to landlord (0,1)	-11.580	5.316
	Number of landlord contacts for maintenance	1.130	2.777
	Number of persons per room	6.970	4.042
Dwelling Unit Features	Area per room (natural log)	19.315	6.003
	Total number of rooms (includes kitchen & bath) (natural log)	53.399	27.081
	Building age (years)	-0 248	4.920
	Dishwasher and/or disposal provided (0,1)	11.055	4.354
	Adequate light and ventilation (0,1)	3.824	3.186
	Average surface and structural quality (4 point scale)	11.639	6.142
	Parking facilities provided (0,1)	8.631	5.287
	Large multifamily structure (0,1)	10.207	5.964
	Recent interior painting or papering (0,1)	7.935	4.274
	Many high quality features (0,1)	14.462	4.825
	Presence of private yard (0,1)	1.807	1.461
Neighborhood Features	Median income of census tract (dollars)	0 002	6.157
	Quality of block face landscaping (4 point scale)	4.450	6.754
	Quality of adult recreation facilities	17.053	8.676
Interaction, Tenure Characteristics	Interaction, length of residence (natural log) x Phoenix	-2.760	3.584
	Interaction, related to landlord x Phoenix (0,1)	-4.123	1.279
	Interaction, number of landlord contacts for maintenance x Phoenix	0.402	0.701
	Interaction, number of persons per room x Phoenix	-0.084	0.032
Interaction, Dwelling Unit Features	Interaction, area per room (natural log) x Phoenix	18.946	4.367
	Interaction, total number of rooms (includes kitchen and bath)(natural log) x Phoenix	24.130	7.924
	Interaction, building age (years) x Phoenix	-0.060	0.789
	Interaction, dishwasher and/or disposal provided x Phoenix (0,1)	-0.021	0
	Interaction, adequate light and ventilation x Phoneix (0,1)	3.981	2.295
	Interaction, average surface and structural quality (4 point scale) x Phoenix	4.258	1.767
	Interaction, parking facilities provided x Phoenix(0,1)	-2.575	1.236
	Interaction, large multifamily structure x Phoenix (0,1)	-3.353	1.348
	Interaction, recent interior painting and papering x Phoenix (0,1)	-4.503	1.927
	Interaction, many high quality features x Phoenix (0,1)	-7.522	2 091
	Interaction, presence of private yard X Phoenix (0,1)	0.170	0.095
Interaction, Neighborhood Features	Interaction, median income of census tract (dollars) x Phoenix	0	0 514
	Interaction, quality of block face landscaping (4 point scale) x Phoenix	-1.488	1.546
	Interaction, quality of adult recreation facilities x Phoenix	-5.266	1.718
CONSTANT		-131.446	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Finally, neighborhood quality is apparently more expensive in Pittsburgh than in Phoenix. Although the coefficients for Census tract median income are not different in the two cities, good block face landscaping and adult recreation facilities--proxy variables for neighborhood quality--both cost more in Pittsburgh.

The impact of these market differences on the use of the common site index is difficult to assess. Clearly, the coefficients in a pooled estimation are weighted averages of somewhat different costs in the two markets. Nevertheless, except for space and tenure, the differences are not necessarily extreme, and tenure does not affect the housing index. On the other hand, the pooling of experimental households or the direct comparison of housing services in Phoenix and Pittsburgh are both useful for certain analytical purposes. The common site index provides an approximate way to compare housing change in a combined sample.

4.2 TESTS FOR MARKET SEGMENTATION WITHIN PITTSBURGH AND PHOENIX

The second major indexing problem dealt with in this chapter is the question of submarket price differences within either Phoenix or Pittsburgh. Empirical research has suggested that the cost gradients for housing attributes may vary substantially within urban areas (Strazheim, 1973; Quigley, 1973; Schnare and Struyk, 1974; King, 1973; Kain and Quigley, 1975). These studies, and others, have used many different stratification schemes to test for such variations, including stratification by tenure type, by size of dwelling unit, by political boundaries, by Census tracts, by central city and suburban areas, and by neighborhood characteristics such as racial composition.

The findings of such research studies on cost variations in housing attributes are relevant both to the analysis of household response to housing allowances and to the derivation and interpretation of the hedonic quality measures used in the Demand Experiment. If submarket differences in housing attribute costs are large and systematic, it would be inappropriate to interpret a single equation estimated for the entire market as reflecting implicit costs across different submarkets.

If households do not have equal access to all submarkets--as is the case with the submarkets formed by racial discrimination--they may face very different opportunities to use a housing allowance. Thus, for example, if minority submarkets contain relatively few units that meet the Minimum Standards requirement, and if these are relatively expensive compared with similar units in the nonminority submarket, then minorities would be expected to be able to meet these standards and to take advantage of the housing allowance offer less often than nonminorities. Similarly, minorities would generally not be able to obtain the same housing quality as nonminorities in meeting Minimum Rent requirements.[1]

The discussion in Appendix I shows that hedonic indices of housing quality will not be perfectly comparable across submarkets in which relative prices differ. Nevertheless, it is important to the analysis to have a single housing index that applies to all participants in a given city. For one thing, since some of the relevant analysis samples are somewhat small, it may not be feasible to use separate indices for different subgroups. Also, there is no consistent way to measure housing change for a household that moves from one submarket area to another area with separate indices. A market-wide index can at least adjust for the average overall differences in the rent/quality relationship across submarkets.

Generally speaking, the stratification of any sample of households depends, at least to some extent, on the purpose of the research and, of course, on the sample size. There is no a priori method of determining the best way to define submarkets. Both the type and degree of segmentation inherent in any given housing market are likely to vary considerably from time to time; moreover, unless the sample is extremely large, a stratification scheme based on small geographic areas or on several attributes of the housing stock is infeasible. A complex stratification scheme is warranted only when research focuses on the variations of individual attribute costs

---

[1] While hedonic regressions can be used to estimate price differences for equivalent units, availability must be investigated separately. This might be done by comparing the types of housing attributes purchased in the two submarkets, while controlling for other factors.

and when the sample is large. Because the Demand Experiment focuses on consistent prediction of overall housing quality rather than on variation in individual attribute costs, another approach to testing for segmentation is more appropriate. The issue then becomes the extent of submarket variation relative to overall variation in estimated attribute costs.

In many respects, because of the nature of the Pittsburgh and Phoenix samples, observation is already limited to portions of the housing market that are relatively homogeneous. Stratification by type of tenure is not applicable because the sample comprises only renter households; similarly, all sample households have low or very moderate incomes. However, tests for the effects of two other potential sources of structural differences in attribute costs are applicable: (1) race of household or racial submarkets, and (2) central city and suburban housing markets. Tests are made for central city and suburban segmentation since the housing stock often differs widely across these areas, and suburban housing may be relatively unavailable to low-income households. Major emphasis, however, is given to exploring the effects of racial segmentation. One reason is that racial segmentation appears to be likely, and many previous studies have found evidence of racial price premiums. Second, such findings would directly reflect on the equity of a housing allowance, that is, the ability of minority households to effectively translate subsidy dollars into improved housing.

The analysis of racial or ethnic submarkets focuses on two issues that are analytically separable--the existence of price discrimination, and the existence and strength of structural differences in the attribute prices of the submarkets. In using hedonic indices to address the second issue, it must be determined whether costs differ by more than a proportional shift and whether the differences appear large enough to have any practical effect. If they do not, then a market-wide hedonic index is appropriate. Thus, price discrimination could exist without destroying the ability to index across racial submarkets. A proportional shift in the quality attribute coefficients could be accommodated by omission of the shift term in calculating the index, thus reducing the different submarkets to a common base. The tests for market segmentation are shown below. The analysis of price discrimination per se is made in Chapter 5.

Tests for Market Segmentation

Separate linear and semilog equations have been estimated for the central city and suburban areas, for the white and minority submarkets, and for the major racial or ethnic groups in each city. The white submarket comprices Census tracts with more than 80 percent white residents. The minority submarket is defined as tracts with more than 50 percent minority residents--that is, black households in Pittsburgh and Spanish American households in Phoenix. Equations for white and black households were estimated in Pittsburgh and for white and Spanish American households in Phoenix. All of these equations are presented in Appendix II.

The summary statistics necessary to assess submarket differences are given in Tables 4-8 through 4-13. Each of these tables describes the separate submarket equations, the full-sample equation relevant to these submarkets, and finally, the full-sample equation with full interaction between attributes and submarket. The last equation is, of course, equivalent to the stratified equations; information concerning the stratified equations is presented in order to show how predictive power varies by submarket. Finally, each table includes the appropriate tests for overall differences by submarket, for proportionate shifts between submarkets, and for comparison of the standard errors.

Tables 4-8 and 4-9 present this information for the central city and suburban submarkets.[1] Attribute cost structures might differ between these areas for a number of reasons. Generally, central city housing stock tends to be older and more dense than suburban housing stock. For example, a strong desire for attributes that are only available in the suburbs, combined with the limited mobility of households that have central workplaces but little access to transportation, could tend to segment these two markets. Also, some attribute combinations might not be available in a continuous mix. For example, because of suburban zoning laws and income distribution, good suburban schools may be associated only with large house lots.

[1]The equations in these tables are based on the regular hedonic variable list described in Chapter 3, Tables 3-2 through 3-5.

Table 4-8

TEST FOR MARKET SEGMENTATION
BY CENTRAL CITY AND SUBURBS: PITTSBURGH

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION			
Central city	813 (k=35)	.65	.63	22.25801	.16914
Suburbs	770 (k=35)	.69	.67	22.33512	.17432
Full sample	1583 (k=35)	.66	.65	45.78433	.17198
Full sample with shift term	1583 (k=36)	.66	.65	45.78339	.17203
Full sample with complete interaction	1583 (k=70)	.67	.66	44.59313	.17168
Submarket Tests:					
F-test for submarket					1.1542 (35,1513)
F-test for proportional shift					1.1882 (34,1513)
Comparison of standard errors					0.0017
		LINEAR EQUATION			
Central city	820 (k=33)	.64	.62	265819.91	18.37834
Suburbs	779 (k=33)	.68	.67	297010.37	19.95338
Full sample	1599 (k=33)	.66	.65	577674.88	19.20640
Full sample with shift term	1599 (k=34)	.66	.65	577640.172	19.21195
Full sample with complete interaction	1599 (k=66)	.66	.65	562830.281	19.16098
Full sample with proportional change	1599 (k=2)	.66	.66	577249.6	19.012
Submarket Tests:					
F-test for submarket					1.2264 (33,1533)
F-test for proportional shift					0.623 (32,1533)
Comparison of standard errors					0.0024

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-9

TEST FOR MARKET SEGMENTATION
BY CENTRAL CITY AND SUBURBS: PHOENIX

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION			
Central city	1217 (k=28)	.80	.80	33.09958	.16685
Suburbs	377 (k=28)	.83	.81	9.19964	.16236
Full sample	1593 (k=28)	.80	.80	43.05891	.16587
Full sample with shift term	1593 (k=29)	.80	.80	43.03664	.16588
Full sample with complete interaction	1593 (k=56)	.81	.80	41.92957	.16517

Submarket Tests:

F-test for submarket	1.4766 (28,1537)
F-test for proportional shift	1.5028 (27,1537)
Comparison of standard errors	0.0042

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		LINEAR EQUATION			
Central city	1217 (k=25)	.79	.78	528785.117	21.06209
Suburbs	377 (k=25)	.81	.80	158964.950	21.25100
Full sample	1593 (k=25)	.79	.78	705454.51	21.21102
Full sample with shift term	1593 (k=26)	.79	.78	705313.08	21.21566
Full sample with complete interaction	1593 (k=50)	.79	.79	685692.93	21.07372
Full sample with proportional change	1593 (k=2)	.79	.79	707432.7	21.080

Submarket Tests:

F-test for submarket	1.7775 (25,1167)
F-test for proportional shift	1.051 (24,1167)
Comparison of standard errors	0.0065

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

The statistics indicate, however, that at least there is little or no segmentation in Pittsburgh and Phoenix for the lower-income portions of the central city and suburban market. None of the F-statistics in Pittsburgh is significant at 0.01. In Phoenix, the test is significant only for the linear equation.[1] In addition, according to the standard error comparison test, there is almost no difference in predictive power (far less than 1 percent between the full-sample and the subsample regressions in each city).

The summary information for the minority submarket and white submarket equations, and the test statistics for market segmentation, are given in Tables 4-10 and 4-11. Recall that the minority submarket is defined as a Census tract in which over 50 percent of the residents are minorities. The white submarket in each city is defined as a tract with less than 20 percent minority residents.[2] Tables 4-12 and 4-13 describe the equations for the white and minority households in each site. Again, black households constitute the dominant minority in Pittsburgh, Spanish Americans in Phoenix.

Again, the subsample equations from which these statistics are derived are presented in Appendix II. With one important exception, the linear and semilog equations in each site are identical to those described in Chapter 3, the final hedonic equations. The exception concerns the derived factor scores used to describe neighborhood quality. In order to assess market segmentation based on the racial composition of the neighborhood, it was necessary to exclude variables describing Census tract racial composition from the principal components analysis. Otherwise, problems of multicollinearity or of biased coefficients due to omitted variables would occur when

---

[1]The critical value of the F-statistic being used is $F_{(.99)}(27,1000) \approx 1.738$. Most tables do not present the critical F-value for samples this large. The value of F cited above was calculated by interpolating for the degrees of freedom in both the numerator and denominator. The value chosen (27,1000) is representative of the sample sizes in the racial submarket equations.

[2]Separate equations were not estimated for the remaining Census tracts, those having 20 to 50 percent minority residents. These mixed neighborhoods are considered in the analysis of price discrimination in Chapter 5, however.

Table 4-10

TEST FOR MARKET SEGMENTATION
BY WHITE AND GHETTO SUBMARKETS: PITTSBURGH

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION			
White submarket	1180 (k=34)	.67	.66	34.62440	.17382
Ghetto submarket	239 (k=34)	.70	.65	5.34945	.16154
Full sample	1419 (k=34)	.66	.66	41.22907	.17253
Full sample with shift term	1419 (k=35)	.67	.66	41.01467	.17215
Full sample with complete interaction	1419 (k=68)	.67	.66	39.97386	.17201
Submarket Tests:					
F-test for submarket				1.2477 (34,1351)	
F-test for proportional shift				1.0644 (33,1351)	
Comparison of standard errors				0.0003	
		LINEAR EQUATION			
White submarket	1192 (k=32)	.67	.66	449203.25	19.67852
Ghetto submarket	241 (k=32)	.67	.62	57674.98	16.61195
Full sample	1433 (k=32)	.66	.65	521628.77	19.29575
Full sample with shift term	1433 (k=33)	.66	.65	519548.30	19.26411
Full sample with complete interaction	1433 (k=64)	.67	.65	506878.23	19.24199
Full sample with proportional change	1433 (k=2)	.66	.66	521118.6	19.083
Submarket tests:					
F-test for submarket				1.2449 (32,1369)	
F-test for proportional shift				0.618 (31,1369)	
Comparison of standard errors				0.0028	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-11

TEST FOR MARKET SEGMENTATION
BY SPANISH AMERICAN AND WHITE SUBMARKETS: PHOENIX

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION			
White submarket	912 (k=27)	.78	.77	22.32001	.15881
Spanish American submarket	214 (k=27)	.73	.69	5.20061	.16677
Full sample	1126 (k=27)	.82	.82	28.76228	.16178
Full sample with shift term	1126 (k=28)	.82	.82	28.75682	.16183
Full sample with complete interaction	1126 (k=54)	.83	.82	27.52062	.16023
Submarket Tests:					
F-test for submarket					1.7906 (27,1072)
F-test for proportional shift					1.8513 (26,1072)
Comparison of standard errors					0.0096
		LINEAR EQUATION			
White submarket	912 (k=25)	.77	.77	401102.14	21.26501
Spanish American submarket	214 (k=25)	.69	.65	49933.82	16.25423
Full sample	1126 (k=25)	.80	.80	490643.83	21.11006
Full sample with shift term	1126 (k=26)	.80	.80	490475.37	21.11603
Full sample with complete interaction	1126 (k=50)	.82	.81	451035.96	20.47385
Full sample with proportional change	1126 (k=2)	.80	.80	488119.00	20.839
Submarket Tests:					
F-test for submarket					3.7789 (25,1076)
F-test for proportional shift					1.838 (24,1076)
Comparison of standard errors					0.0301

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-12

TEST FOR MARKET SEGMENTATION
BY WHITE AND BLACK HOUSEHOLDS: PITTSBURGH

EQUATIONS	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION			
White households	1205 (k=34)	.67	.66	35.80872	.17487
Black households	362 (k=34)	.67	.64	7.95971	.15578
Full sample	1567 (k=34)	.66	.65	45.53771	.17235
Full sample with shift term	1567 (k=35)	.66	.65	45.53755	.17241
Full sample with complete interaction	1567 (k=68)	.67	.66	43.76843	.17088
Submarket Tests:					
F-test for submarket					1.7812 (34,1499)
F-test for proportional shift					1.835 (33,1499)
Comparison of standard errors					0.0085
		LINEAR EQUATION			
White households	1219 (k=32)	.66	.65	458821.95	19.66060
Black households	364 (k=32)	.65	.62	91475.78	16.59908
Full sample	1583 (k=32)	.65	.65	571369.75	19.19344
Full sample with shift term	1583 (k=33)	.65	.65	571321.39	19.19881
Full sample with complete interaction	1583 (k=64)	.67	.65	550297.72	19.03356
Full sample with proportional shift	1583 (k=2)	.65	.65	570787.5	19.001
Submarket Tests:					
F-test for submarket					1.8151 (32,1519)
F-test for proportional shift					0.907 (31,1519)
Comparison of standard errors					0.0083

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-13

TEST FOR MARKET SEGMENTATION
BY WHITE AND SPANISH AMERICAN HOUSEHOLDS: PHOENIX

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		SEMILOG EQUATION			
White households	1065 (k=27)	.80	.79	26.37572	.15941
Spanish American households	378 (k=27)	.79	.78	10.4100	.17222
Full sample	1443 (k=27)	.81	.80	37.90952	.16362
Full sample with shift term	1443 (k=28)	.81	.80	37.84333	.16354
Full sample with complete interaction	1443 (k=54)	.81	.81	36.78572	.16274

Submarket Tests:

F-test for submarket	1.5691 (27,1389)
F-test for proportional shift	1.5386 (26,1389)
Comparison of standard errors	0.0054

EQUATION	NUMBER OF OBSERVATIONS	R^2	$\bar{R}^2$	RESIDUAL SUM OF SQUARES	STANDARD ERROR
		LINEAR EQUATION			
White households	1065 (k=25)	.78	.78	462440.78	21.08683
Spanish American households	378 (k=25)	.78	.76	149191.66	20.55819
Full sample	1443 (k=25)	.79	.79	631218.43	21.09851
Full sample with shift term	1443 (k=26)	.79	.79	629971.12	21.08509
Full sample with complete interaction	1443 (k=50)	.80	.79	611632.44	20.95413
Full sample with proportional change	1443 (k=2)	.79	.79	630214.1	20.913

Submarket Tests:

F-test for submarket	1.7830 (25,1393)
F-test for proportional shift	0.871 (24,1393)
Comparison of standard errors	0.0068

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

variables that represent minority submarkets or households were entered in the equations or used to stratify them. Thus, the principal components analysis based on Census variables was reestimated, excluding the variables for percentage black and percentage Spanish American in the Census tract. The solutions are very similar to the original group, particularly in Phoenix, where minority status did not appear to constitute an independent dimension of neighborhood. The factor coefficient matrices for these modified factor scores is given in Appendix VII. The revised factor scores were then entered in the hedonic equations and retained or excluded according to the test level used in the final equations--that is, $t \geq 1.0$.

The results indicate that while some degree of market segmentation exists in some of these submarkets, it is extremely small. In Phoenix, the F-statistics are significant for minority and white submarkets, and, in the linear equation, for minority and white households. In Pittsburgh, the F-tests are significant for households but not for submarkets. This discrepancy is difficult to explain. When the shift term for ghetto submarket is entered in both the linear and semilog equations, it is highly significant (refer to Appendix II and Chapter 5).

In any event, the finding of submarket differences appears to be of extremely limited practical significance. There is very little change in $\bar{R}^2$ for any of these equations. Except for the linear equation in Table 4-11 (Spanish American and white submarkets), the standard errors of the full-sample equations, as compared with the subsample equations, fall by less than 1 percent. This would seem to be operationally irrelevant.

4.3 INDEXING THE CHANGE IN HOUSING

Evaluation of the change in housing services over the two-year experimental period is done with reference to the weights (coefficients) estimated at the baseline period. As described in Chapter 2, the housing index, derived from the hedonic equation estimated at baseline, is

$$(1) \qquad \hat{Q}_j = \alpha + \Sigma X_{ij}\hat{\beta}_i$$

and the change in housing is indexed as

$$(2) \qquad \Delta\hat{Q}_j = \Sigma \Delta X_{ij}\hat{\beta}_i$$

where X_{ij} (ΔX_{ij}) is the amount of (change in) the i^{th} attribute in the dwelling unit and neighborhood, and the $\hat{\beta}_i$ are the estimated hedonic coefficients.[1]

There are good reasons for using the pre-experimental sample to specify the equation. First, the sample of baseline households is by far the largest available for specification of the model. One of the uses of the housing index is to investigate the extent to which experimental households spend more or less than normal to obtain a given quality of housing. Thus, for time periods after the experiment began only Control households could be used in estimation. This would reduce the sample size by a factor of about seven in Phoenix and five in Pittsburgh. In addition, since the experimental period is only two years, the attribute cost structure is not expected to shift in a major way. In order to evaluate how well the baseline model predicts rent at two years, this section examines the association among actual rent at two years, rent predicted by the baseline hedonic model and rent predicted by an hedonic equation estimated using Control households at the two-year cross section.

Like the other indexing issues discussed in this chapter, the salient issue is whether relative attribute prices have shifted. The price of housing is expected to increase over the experimental period because of inflation. If inflation merely results in proportionate shifts in all attribute prices, however, using baseline weights presents no problem.[2]

The conclusion reached is clear: the baseline hedonic model is quite stable over the two-year period and provides very reasonable estimates of rent and housing services. Since the baseline model predicts rent as efficiently as an equation estimated using the two-year sample, the baseline weights are

---

[1] Strictly speaking, the hedonic equations have been specified on the enrollment sample. The term baseline is also used in this section and is meant to indicate the pre-experimental (enrollment) sample.

[2] For a discussion of the usual Laspeyres-Passche indexing problem in the context of using hedonic indices over time and some suggested alternative weighting schemes see Griliches (1971) and the references cited there. Since the end period sample is quite small (Control households only), the benefit to be gained from either weighting schemes for coefficients or pooled cross-section and time-series estimation are probably quite small.

appropriate for indexing the change in housing. The efficiency of the estimated baseline model is assessed by comparing two estimates for rent at two years--one using the baseline weights and allowing only for a proportional shift in costs due to inflation and one that estimates a completely new set of weights at two years. The sample for this comparison is Control households at two years.

Summary statistics for the linear and semilog equations estimated using this sample are presented in Table 4-14, which also presents summary statistics for a baseline equation estimated for the same group of Control households.[1] Since the hedonic model is specified on the baseline sample, a fall in explanatory power at the two-year period would be expected. While there is some drop in Phoenix, the explanatory power in Pittsburgh is almost as high at two years as at baseline.[2]

The test statistics are derived as follows. For the semilog model, a simple proportional inflation is represented by

$$\ln R_{t_0} = X_{t_0}\beta_{t_0} + \varepsilon_0 \tag{3}$$

$$\ln R_{t_3} = X_{t_3}\beta_{t_0} + \alpha + \varepsilon_3$$

$$\ln R_{t_3} - X_{t_3}\beta_{t_0} = \alpha + \varepsilon_3,$$

where subscripts refer to the time periods. Thus, proportional inflation is tested by comparing the sum of squared residuals from

$$\ln R_{t_3} - \widehat{\ln R}_{t_{3(B)}} = \alpha + \varepsilon \tag{4}$$

with the squared residuals from an equation estimated using Control households at two years

---

[1] The complete equations are listed in Appendix II.

[2] This may be partly explained by the difference in mobility rates by sites; in Pittsburgh, a greater proportion of the two-year sample consists of the same dwelling units as in the baseline sample because fewer households moved during the two-year period.

Table 4-14

DESCRIPTIVE STATISTICS FOR HEDONIC REGRESSIONS
ESTIMATED AT ENROLLMENT AND AT TWO YEARS
(Control Households Only)

STATISTIC	SEMILOG EQUATION		LINEAR EQUATION	
	ENROLLMENT	TWO YEARS	ENROLLMENT	TWO YEARS
PITTSBURGH				
Sample	319	321	324	322
R^2 and $\bar{R}^2$	.70 - .67	.69 - .65	.70 - .67	.67 - .63
Residual sum of squares	8.63511	9.85904	117569.21	169527.77
Standard error	.17437	.18567	20.100	24.2199
PHOENIX				
Sample	250	241	250	241
R^2 and $\bar{R}^2$	.84 - .82	.67 - .63	.84 - .82	.70 - .67
Residual sum of squares	5.97506	12.7236	85834.64	192399.97
Standard error	.16406	.2444	19.48844	29.77643

SAMPLE: All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

(5) $$\ln R_{t_3} = X_{t_3}\beta_{t_3} + \varepsilon,$$

where

$\ln R_{t_3}$ = log of actual rent at t_3

$\widehat{\ln R}_{t_{3(B)}}$ = predicted log of rent at t_3 using baseline coefficients (B).

The test statistic for the semilog equation is

(6) $$F = \left[\frac{SSR(4)}{SSR(5)} - 1\right]\left[\frac{n-k}{k-1}\right]$$

where

$SSR(\cdot)$ = the sum of the squared residuals from Equation $(\cdot)$

n = the number of observations at t_3

k = the number of parameters in the unrestricted equation.

For the linear model, a proportional inflation in weights is represented by

(7) $$R_{t_o} = X_{t_0}\beta_{t_0} + \varepsilon_0$$

$$R_{t_3} = \alpha X_{t_3}\beta_{t_0} + \varepsilon_3.$$

Thus, proportional inflation is tested by comparing the squared residuals from

(8) $$R_{t_3} = \alpha\hat{R}_{t_{3(B)}} + \varepsilon$$

with the residuals from

(9) $$R_{t_3} = X_{t_3}\beta_{t_3} + \varepsilon.$$

The test statistic for the linear equation is, following the notation of Equation (6),

(10) $$F = \left[\frac{SSR(8)}{SSR(9)} - 1\right]\left[\frac{n-k}{k-1}\right].$$

For the linear model this tests the hypothesis that

(11) $$\beta_{(t_3)} = \alpha\hat{\beta}_{(B)}.$$

If the sums of squared residuals are not significantly different, then the hypothesis that the baseline and two-year models differ by no more than a proportionate shift is not rejected.

Table 4-15 lists the F-statistics for the linear and semilog equations in Pittsburgh and Phoenix; none is significant.[1] Thus, the baseline weights can be used to index housing quality over the two years.

The hedonic estimates of rent at two years can be used to provide an estimate of inflation during the experimental period. The mean difference between actual rent at two years and predicted rent using the baseline coefficients is $\hat{\alpha}$, expressed as

$$\text{(12)} \qquad \ln R_{(t_3)} - \widehat{\ln R}_{t_{3(B)}} = \alpha + \varepsilon .$$

Thus, $\hat{\alpha}$ is an estimate of $\ln(1+\rho)$ where ρ is the rate of inflation from enrollment to two years.

This estimate may be compared with another estimate of inflation: the difference between actual rent at two years and at enrollment for Control households that did not move. As discussed in Chapter 3, a significant discount from rent is associated with length of residence. Since this discount must be adjusted for when comparing the change in rent for non-movers, the estimate is

$$\text{(13)} \qquad \ln\tilde{R}_{t_3} - \ln\tilde{R}_{(t_0)} = \left[\ln R_{(t_3)} - T_{(t_3)}\hat{\beta}_{(B)}\right] - \left[\ln R_{(t_0)} - T_{(t_0)}\hat{\beta}_{(B)}\right]$$

where T represents the length of tenure at enrollment or at two years and $\hat{\beta}_{(B)}$ is the estimated (baseline) coefficient for the length of tenure variable.

As above, this difference estimates $\ln(1+\rho)$. The inflation estimates for Pittsburgh and Phoenix are given in Table 4-16. For both approaches, the

---

[1] Another way of assessing whether the coefficients are the same or whether only a proportionate shift in attribute prices has occurred, is to estimate a combined time-series/cross-section regression. Then, following the approach used in previous sections of this chapter, a dummy variable for shift over time and a complete set of time interaction terms would be included. Since the emphasis here is on the ability to predict rent at two years using baseline weights, an alternative approach is used.

Table 4-15

ASSESSMENT OF THE PREDICTIVE POWER
OF THE BASELINE REGRESSION AT TWO YEARS (CONTROL HOUSEHOLDS)

SEMILOG EQUATION

EQUATION	NUMBER OF OBSERVATIONS	RESIDUAL SUM OF SQUARES	STANDARD ERROR
	PITTSBURGH		
$\ln R_{t_3} - \widehat{\ln R}_{t_{3(B)}} = \alpha + \varepsilon$	321 (k=35)	10.716	.18567
$\ln R_{t_3} = X_{t_3}\hat{\beta}_{t_3} + \varepsilon$	321 (k=35)	9.8590	.18400
	PHOENIX		
$\ln R_{t_3} - \widehat{\ln R}_{t_{3(B)}} = \alpha + \varepsilon$	241 (k=28)	14.540	.24800
$\ln R_{t_3} - X_{t_3}\hat{\beta}_{t_3} + \varepsilon$	241 (k=28)	12.724	.24400
F-test for Pittsburgh		.321	.732 (34,286)
F-test for Phoenix		.490	1.128 (27,213)

LINEAR EQUATION

EQUATION	NUMBER OF OBSERVATIONS	RESIDUAL SUM OF SQUARES	
	PITTSBURGH		
$R_{t_3} = \alpha\hat{R}_{t_{3(B)}} + \varepsilon$	321 (k=33)	182643.40	24.228
$R_{t_3} = X_{t_3}\hat{\beta}_{t_3} + \varepsilon$	321 (k=33)	169527.76	24.220
	PHOENIX		
$R_{t_3} = \alpha\hat{R}_{t_{3(B)}} + \varepsilon$	241 (k=24)	218262.80	30.157
$R_{t_3} = X_{t_3}\hat{\beta}_{t_3} + \varepsilon$	241 (k=24)	192399.97	29.776
F-test for Pittsburgh		.307	.693 (32,288)
F-test for Phoenix		.562	1.264 (23,217)

SAMPLE: All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table 4-16

ESTIMATED RATES OF INFLATION BETWEEN ENROLLMENT AND TWO YEARS

NONMOVER CONTROL HOUSEHOLDS

ESTIMATE	PITTSBURGH	PHOENIX
Mean difference, rent at two years and predicted rent using the Baseline hedonic coefficient	13.0%	7.3%
Mean difference in rent between two years and enrollment (adjusted for tenure discount)	14.8	10.0

SAMPLE: All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

sample is Control households that did not move between enrollment and two years.[1]

The closeness of the two inflation estimates with each site again suggests the reasonableness of the predictions made using the baseline hedonic model. The small difference between the two estimated rates within each site is apparently due to a slight increase in the housing index for nonmover households over the two years. Table 4-17 shows the mean housing index for nonmover Control households at enrollment and two years; the index increases by less than 2 percent in Pittsburgh and by about 3 percent in Phoenix.

The estimated inflation rates in Pittsburgh, 13 to 15 percent, are extremely close to the estimated change in rent component of the national consumer price index. This index rose exactly 13 percentage points between 1973 and 1975.[2] In addition, the Bureau of Labor Statistics rental budget (from the Intermediate Budget figures) rose by 16.8 percent in Pittsburgh between 1973 and 1975. No comparable figure exists for Phoenix. However, to the extent that the inflation rate is likely to be somewhat lower in an expanding and looser housing market, the Phoenix estimates appear reasonable.

[1] In addition, it should be noted that actual rent at two years in Equations (12) and (13) has been adjusted for utilities using the updated estimates of utility costs, since utility cost increases are a likely contribution to inflation. Refer to Appendix VI for a discussion of adjusted rent and original and updated utility adjustment tables. If rent is adjusted using the original utility tables, the difference in rent (adjusted for tenure) indicates an inflation rate of 10 percent in Pittsburgh and 7 percent in Phoenix.

[2] Based on 1967=100, the CPI rent component was 124.3 in 1973 and 137.3 in 1975 (Bureau of Labor Statistics, 1976).

Table 4-17

HOUSING INDEX AT ENROLLMENT AND TWO YEARS
FOR CONTROL HOUSEHOLDS THAT DO NOT MOVE

	MEAN HOUSING INDEX (STANDARD ERROR)	
	ENROLLMENT	TWO YEARS
Pittsburgh N = 185	114.59 (24.59)	116.56 (25.83)
Phoenix N = 122	130.02 (38.03)	134.10 (37.02)

SAMPLE: All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

REFERENCES

Griliches, Zvi, Price Indexes and Quality Change, Cambridge, Mass., Harvard University Press, 1971.

Johnston, J., Econometric Methods, New York, McGraw-Hill, 1972.

Kain, John and John Quigley, Housing Markets and Racial Discrimination, New York, National Bureau of Economic Research, 1975.

Kelly, T. L., "An Unbiased Correlation Measure," Proceedings of the National Academy of Science, vol. 21, 1935, pp. 554-559.

King, Thomas, Property Taxes, Amenities, and Residential Land Values, Cambridge, Mass., Ballinger Publishing Co., 1973.

Ohta, Makoto and Zvi Griliches, "Automobile Prices Revisited: Extensions of the Hedonic Hypothesis," Household Production and Consumption, edited by Nestor E. Terleckyj, New York, National Bureau of Economic Research, 1975.

Quigley, John, "Housing Demand in the Short Run: An Analysis of Polytomous Choice," (paper prepared for the Winter Meetings of the Econometric Society, December 1973, New York), Yale University, Institute for Social and Policy Studies, 1973.

Schnare, Ann and Raymond Struyk, "Segmentation in Urban Housing Markets," (paper presented at the Committee on Urban Economics, Conference on Housing Economics, October 1974, St. Louis, Missouri, Washington University), Washington, D.C., Urban Institute, 1974.

Straszheim, Mahlon, "Housing Market Discrimination and Black Housing Consumption," Quarterly Journal of Economics, 1973, pp. 19-43.

CHAPTER 5

THE EFFECTS OF SEGREGATION ON HOUSING PRICE

Racial segregation has a potentially pervasive effect on the housing market. Racial barriers to locational choice may affect the quality of minority housing, the ability of minorities to become homeowners, and their access to better neighborhoods, better schools, and employment opportunities.[1] Segregation and discrimination may also affect the relative price of minority housing--that is, minorities that pay the same rent as nonminorities may get much less housing quality.

These issues are of obvious importance in the evaluation of a housing allowance. If price discrimination exists, then minorities may be less able than nonminorities to translate a given rental expenditure or allowance payment into improved housing. Even when no price markup exists, limited access to certain types of housing, particularly to higher quality units, may lower the chances for minority households to meet target levels of housing quality.[2]

Most analyses of the price effects of discrimination in the housing market conclude that minorities do pay more than whites for equivalent housing.[3] The estimated discriminatory premiums range from 4 percent to over 30 percent. Other analyses reach the opposite conclusion--namely, that

---

[1] The effects of housing market segregation have been extensively studied, and a lengthy bibliography exists. For some examples, see Myrdal, 1962; Kain, 1962, 1972; Lieberson, 1963; McEntire, 1960; Taueber and Taueber, 1969; Quigley, 1974. Sociologists, demographers, land use planners, and economists have all contributed to research on discrimination. The discussion in the present study is drawn largely from the work of economists, primarily because the present approach to analyzing price effects is based on hedonic regressions, a method most frequently used by economists.

[2] Previous analyses of housing allowance data have indicated that minority households in both Pittsburgh and Phoenix are less likely to meet Program Minimum Standards than similar white households paying the same amount of rent (Abt Associates, 1975).

[3] See Ridker and Henning, 1967; Haugens and Heins, 1969; Kain and Quigley, 1975; King and Mieszkowski, 1973; Quigley, 1974; Stengel, 1973; Schnare, 1974; Gillingham, 1973; Muth, 1969; Rapkin, 1966.

whites pay more than minorities for equivalent housing.[1] The quality of these analyses of price discrimination varies tremendously. Many studies fail to specify what hypotheses are being tested or what factors might be responsible for observed price differences. Both conceptual and empirical problems complicate interpretation of the results. An even greater problem has been a lack of housing data adequate to define equivalent housing bundles and a failure to appreciate the implications of this problem. For example, several of the analyses lack neighborhood variables in the hedonic equations, and many others rely solely on Census data to describe dwelling units. If some relevant attributes are not controlled for, however, the estimated price differentials for racial groups may reflect only systematic differences in the housing stock or in the neighborhoods of these groups. Finally, the majority of the analyses rely on the full-sample approach, with shift terms to test for price differentials. This approach may be insufficient when relative attribute cost factors vary significantly in racial submarkets. Both a full-sample and a subsample approach are presented below.

Chapter 4 focused on issues that affect the use of hedonic indices in the analysis and presented tests for market segmentation between minority and white neighborhoods and households. Although the extent of segmentation does not appear to be large, various racial or ethnic submarkets in each site were found to differ. This chapter focuses explicitly on the question of whether minorities pay more, less, or the same as whites for equivalent housing.

Two types of price discrimination can be distinguished. Either prices for housing in minority neighborhoods are higher than prices for equivalent housing in white neighborhoods, or minority households pay more for equivalent housing regardless of their neighborhood. This study is concerned primarily with price markups within ghetto neighborhoods. Some attempt is also made to assess "pure" price discrimination, which occurs when different prices are charged to minorities and whites for equivalent housing in the same neighborhood, but sample size limits an extensive analysis of "pure" price discrimination.

---

[1] See Bailey, 1966; Lapham, 1971; Daniels, 1975; Berry and Bednary, 1975.

The next subsections discuss the theoretical evidence concerning the effects of discrimination and segregation on housing price, as well as the results of tests for price discrimination that use both the full-sample and submarket equations.

The test results suggest that a price markup for ghetto households exists in Pittsburgh; however, it is only about 4 percent. Discrimination appears definitely to be a function of neighborhood rather than of race in Pittsburgh. Overall, black and white households appear to pay about the same for equivalent housing.

The test results show that no discriminatory markups exist in Phoenix. If anything, black households pay much less than either white or Spanish American households for equivalent housing. In addition, Spanish American households may pay slightly less than white households. No evidence of neighborhood price differences for largely Spanish American Census tracts was found.

Theories of Price Effects in Segregated Markets

Different hypotheses that attempt to explain the effects of racial discrimination predict quite different outcomes. Gary Becker (1957) pointed out that the existence of segregated housing markets is not necessarily reflected in rents. Becker therefore thought that prices would equalize across submarkets. The "equilibrium" theories of Martin Bailey (1966), Richard Muth (1969), and Anthony Pascal (1970) maintain that the price of comparable housing will be higher in the white housing submarket. In contrast, the market separation model advanced by Robert Haugens and James Heins (1969) and John Kain and John Quigley (1975) maintains that constraints on the supply of housing to minorities will lead to a positive discriminatory markup against black households. These hypotheses differ in their assumptions about the degree of equilibrium in the housing market and about the discriminatory mechanisms that might create and maintain price differentials.

Almost all of the analyses have focused on black households and have assumed the presence of some type of racial discrimination against blacks. Socioeconomic differences or black preferences for self-segregation are generally not considered to account fully for the extent of black-white housing segregation (Pascal, 1970). Less analysis has been done concerning the segregation of Spanish American households. However, widespread aversion or attraction to many kinds of "neighbors," whether based on race, ethnicity, or socioeconomic status, may lead to neighborhood externalities in the price of housing (Schnare, 1974).

A representative statement of the equilibrium hypothesis is made by Richard Muth in Cities and Housing (1969). Drawing on earlier analyses, Muth formulates a model of a racist housing market and tests some of his hypotheses in an analysis of Chicago data. Like Becker and Pascal, Muth feels that the most reasonable explanation for residential segregation is that whites have a greater aversion to living among blacks than do blacks. In his analysis, explicit widespread collusion is not a prerequisite for racially based housing premiums. Whites are assumed to bid more than blacks for housing in predominantly white areas, thus leading to two racially separate markets with a common border. The price of housing along the border differs from prices in the interiors of the two areas. If whites have an aversion to living among blacks, they will rent units on the white side of the border only if housing is cheaper there than in the white interior. Thus, blacks may be able to outbid whites along the border. If the housing market is competitive, white owners have an incentive to sell to blacks and the boundary will tend to shift toward the white area. If no external forces prevent this movement, and if black and white demands for housing are changing at the same rate, then the black housing supply will grow, relative to the supply for whites, and housing prices for blacks will fall, relative to those for whites.

In equilibrium, prices for boundary housing are equal for both groups; prices in the white interior are higher than prices at the white boundary; and prices in the black interior are lower than prices in the white interior. Muth admits, however, that if black demand for housing is growing rapidly, relative to white demand, and if discriminatory mechanisms slow down the expansion of black submarkets, then prices in black areas will be higher than they otherwise would be.

In summary, the equilibrium model of price structure in racial submarkets focuses on the price advantages that would exist for minorities under long-run equilibrium. The model depends on particular assumptions about the preferences of racial groups--namely, that whites have an aversion to living with blacks, and that blacks either prefer integration or are indifferent to housing location. In general, the ghetto is assumed to be capable of expansion; housing suppliers rent or sell to the highest bidder, and collusive activity or other barriers (such as restrictive zoning or red-lining) are assumed not to inhibit the expansion of ghetto boundaries.

Alternative market separation models, which hypothesize higher prices for equivalent housing in segregated markets, challenge many of Muth's assumptions. First, these models suggest that disequilibrium conditions are likely to persist for long periods, and they focus on the "disequilibrium" adjustment process rather than on long-run equilibrium. Second, they posit discriminatory mechanisms based on collusive activity and overt restrictions (some of which were legal until recently), rather than on the preferences of individual consumers acting through a competitive market.

One of the first market separation models was developed by Haugens and Heins (1969). They contend that blacks will tend to pay higher prices for equivalent housing, due to the containment of black submarkets. The extent of the price differential in this model will depend on several characteristics of the metropolitan area--the rate of increase in the black population, the rate at which whites pull back from ghetto borders, and the extent of spillover opportunities for blacks to move to relatively unpopulated areas of the city. Thus, in a city with a highly centralized ghetto housing market and with strong increases in black demand, ghetto prices will be pushed well above white market prices.

The most extensive analysis of the effects of market separation and discrimination in the housing market is that of John Kain and John Quigley (1975). They hypothesize that price premiums for blacks arise because of constraints on the supply of housing to blacks and because of costs associated with the transfer from white to black occupancy. Ghetto housing is expected to be more expensive than equivalent housing in white areas. Within each submarket, however, whites and blacks are expected to pay the same amount.

Kain and Quigley's analysis of the effects of discrimination is based on their theory of the urban housing market, which introduces many factors not considered in previous models. In particular, they emphasize the importance of the existing stock in the market, the heterogeneity of the stock, and the spatial variation in the cost of housing bundle attributes. Kain and Quigley point out that if the conventional long-run equilibrium model of the housing market were relevant, then the residential segregation that arises from white prejudices would result in higher prices for whites, as predicted by the Bailey, Muth, and Pascal models. However, they reject many of the tenets of conventional equilibrium analysis and claim that the impact of discrimination must be considered in terms of the characteristics of ghetto housing and the possible variation in attribute prices across racial submarkets. Furthermore, they point out that housing market segregation modifies the traditional concepts of residential choice. Since blacks are largely limited to certain residential areas, this creates a situation in which location rents for equally accessible housing are not necessarily the same in white and black submarkets.

The discriminatory mechanism in the Kain-Quigley model recognizes the role of collusive behavior--including deed restrictions, appraisal practices, the actions of real estate brokers, discrimination in financing, and zoning restrictions. Although some of these practices are now illegal, Kain and Quigley consider that their effects are likely to last a long time.

The role of the existing housing stock in the total supply of housing is extremely important in black housing markets. The supply of housing to blacks is most likely to increase through the conversion of existing units or the expansion of ghetto boundaries, rather than through new construction. Also, the factors that affect a change in housing quality are likely to differ in black and white housing markets. In particular, problems such as abandoned buildings, crime, and undesirable land use in ghetto neighborhoods can lead to negative external economies which affect the value of and return to structures and thus the desirability of maintaining ghetto properties. To the extent that conversion and new construction are limited, increased supply occurs only through expansion of ghetto boundaries.

Kain and Quigley consider that a premium is necessary to shift housing to the ghetto submarket. The markup may be a constant amount, as in Equation (1), or proportional to value, as in Equation (2):

$$R_k^B = R_k^W + \alpha \tag{1}$$

$$R_k^B = R_k^W (1 + \beta) \tag{2}$$

where

R_k^B = monthly rent (or market value) for housing type k in the ghetto submarket

R_k^W = monthly rent for housing type k in the white submarket

α, β = discriminatory premiums.

It is important to note that the premium will vary across different metropolitan areas and across the same area over time, as a function of the ease with which the black housing supply can be expanded at the ghetto boundaries, the extent of racial prejudice, the degree of market organization, and the types of discriminatory mechanisms available. When excess demand within the ghetto raises ghetto housing prices relative to white submarket housing prices (plus the premium), then housing units shift from white to black occupancy, often at the ghetto boundaries.

The importance of boundary areas outlined in the Bailey-Muth models, and the effects hypothesized by the market separation model have been thoroughly analyzed by Thomas King and Peter Mieszkowski (1973). They distinguish three types of potential price effects. First, rent differences between ghetto and white submarkets, with ghetto units bearing a discriminatory price premium, arise from the "funneling effects of market separation." This type of price premium is attached to submarkets, independent of the race of the occupant. Second, and by contrast, "pure racial price discrimination" occurs when blacks pay more than whites for equivalent housing in the same neighborhood or submarket. If whites have an aversion to living with blacks, this is most likely to occur in racially mixed submarkets. Third, "white tastes for segregation" are measured by differences in rent between the white interior and the white boundary area. If whites prefer to live away from blacks, white boundary rents will be lower.

The tests of racial price discrimination in terms of racial submarkets and pure racial discrimination are presented below.

Full-Sample Tests for Price Discrimination

Several different approaches are used to test the hypotheses concerning price effects. Since no a priori presumptions have been made concerning supply conditions, levels of prejudice, or discriminatory mechanisms in the Pittsburgh and Phoenix housing markets, the null hypothesis assumes that no price differentials exist. Specification of the appropriate racial variables was guided by comparability with previous analyses, and a distinction was made between race of household and the racial composition of submarkets, such as that made by King and Mieszkowski (1973).

Pittsburgh. Four separate (full-sample) hedonic equations have been specified for Pittsburgh; each includes a different definition of the race of household on racial submarket variables used to test for price effects, but is identical in all other respects. The equations used to make these tests are given in Appendix II. With one exception, the equations are similar to the final full-market regressions presented in Chapter 3.[1]

The equations are estimated in both semilog and linear form. Table 5-1 lists the estimated coefficients for the race or submarket variables included in each semilog and linear equation in Pittsburgh.

First, following the majority of previous studies, the simplest test of a "race" effect is to include a dummy variable that represents black heads of household (Equation (1) in Table 5-1). This specification tests the general hypothesis that black households, no matter where they live, pay a different price than white households for comparable housing. As seen in the table, the coefficient for the variable that represents black households is very small and not statistically different from zero. Thus,

---

[1]The exception is that the derived factor scores based on Census variables do not include the variables "percentage black" or "percentage Spanish American" in the Census tract. These modified factors are listed in Appendix VII. The equations used in Chapter 5 are exactly like those used to test for racial submarkets in Chapter 4, as discussed in Section 4.2.

Table 5-1

RESULTS OF TESTS FOR PRICE DISCRIMINATION: PITTSBURGH

EQUATION NUMBER	RACE OR SUBMARKET VARIABLE INCLUDED IN THE EQUATION	SEMILOG EQUATION		LINEAR EQUATION	
		COEFFICIENT	t-STATISTIC	COEFFICIENT	t-STATISTIC
1	Black head of household	-0.001	0.063	-0.519	0.390
2	Ghetto submarket (Census tracts > 50% black)	0.040**	2.773	3.964**	2.487
	Racially mixed submarket (Census tracts 20 to 50% black)	0.011	0.751	1.082	0.654
3	Ghetto submarket	0.037*	2.329	3.550*	2.045
	Racially mixed submarkets in the central city	0.017	0.769	0.650	0.261
	Racially mixed submarkets in the suburbs	0.000	0.000	0.724	0.321
	White submarket in the central city	-0.006	0.526	-0.790	0.6
4	Black household in the ghetto submarket	0.031*	2.024	2.882+	1.700
	White household in the ghetto submarket	0.043	1.230	4.102	1.079
	Black household in the mixed submarket	0.004	0.202	0.112	0.045
	White household in the mixed submarket	0.005	0.266	0.510	0.226
	Black household in the white submarket	-0.070**	3.186	-8.230**	3.381

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a C* neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

NOTE: The equations used to test for discrimination are exactly like those used to test for racial submarkets in Chapter 4. The derived factor scores (for Census tracts) do not include the variables percentage black or percentage Spanish American in the Census tract (see Appendix VII). The remaining independent variables are those listed in the final equations in Chapter 3.

* Significant at the 0.05 level.

** Significant at the 0.01 level.

\+ Significant at the 0.10 level.

no price differential appears to exist, on the average, for this sample of black households.

The above approach does not adequately identify minority submarkets, however. Many black households do not live in predominantly black neighborhoods, and both the equilibrium and market separation models emphasize price differentials as a function of submarkets rather than of race of household per se. Equation (2) in Table 5-1 presents an alternative definition of racial submarkets, using a dummy variable that represents residence in "ghetto" submarkets or in racially mixed submarkets. "Ghetto area" represents all Census tracts in which black households exceed 50 percent of the population. "Mixed submarket" represents all Census tracts in which black households comprise 20 to 50 percent of the population. Finally, "white submarket" represents Census tracts with fewer than 20 percent black households--this variable is the excluded category in Equation (2).

This submarket approach shows evidence of a neighborhood price markup in Pittsburgh. The semilog specification of Equation (2) indicates that housing in all black areas is 4 percent (or $4.00, based on the linear equation) more expensive than in all white areas.

The static nature of the analysis should be kept in mind when interpreting the present results. The dynamic effects of the transition from white to minority occupancy, especially in expanding areas of the ghetto boundary, may have an important influence on price effects of segregation. This issue has been analyzed in many studies of owner-occupied units.[1] Many of these studies draw similar conclusions. For example, many studies show that, following a temporary drop in prices prior to changeover--the "expectations" effect--housing values rise rapidly after black entry. No attempt has been made to adequately identify "transition" areas in the present study. As a crude proxy for potential transition areas, however, racially mixed markets in the central city have been distinguished from racially mixed markets in the suburbs [Equation (3)]. Central city mixed areas approximate potential spillover areas in Pittsburgh, which has no major suburban ghetto areas. Equation (3) indicates that no price

[1] See Laurenti, 1972; and the bibliography cited in Kain and Quigley, 1975.

difference exists for housing in central city racially mixed markets, relative to housing in white suburban submarkets (the excluded category). Also, no premium is evident for mixed markets in suburban areas.

Overall, the results of Equations (1) through (3) in Pittsburgh provide partial support for the hypotheses of the market separation model, in showing that price effects are not a simple function of race of household. Housing in black neighborhoods is generally more expensive than housing in primarily white submarkets; however, housing in central city racially mixed markets is no more expensive than housing in white areas. These results therefore cannot confirm Kain and Quigley's hypothesis that a premium is necessary to shift housing to the ghetto market. It should be noted, however, that if accessibility or other characteristics of the central city area relative to those of suburban areas have not been adequately controlled for, the results could reflect factors other than racial composition.

An additional test for another type of price effect has been made. "Pure" price discrimination occurs when minority households pay more than white households that live in the same submarket for comparable housing. Thus the dummy variables included in Equation (4) distinguish black and white households that live in primarily ghetto areas, mixed areas, and primarily white areas (white households living in the white submarket comprise the excluded category). These distinctions are possible in the present sample because an adequate sample of black households live in white submarkets, and a small, but adequate, sample of white households live in black neighborhoods.

The test shows no evidence of pure price discrimination. Both blacks and whites pay more for housing in the ghetto market. No significant effects are seen in the mixed market. In the white submarket, however, blacks pay substantially less than whites, a result that is compatible with neither the equilibrium nor the market segmentation model; this may in part reflect the inaccuracy of using Census tracts as a data base, or the mechanisms involved may not be totally racial.

<u>Phoenix</u>. Table 5-2 presents the results of a similar analysis in Phoenix; five separate equations, each with a different specification of racial, ethnic, or submarket variables, have been estimated in both semilog and

Table 5-2

RESULTS OF TESTS FOR PRICE DISCRIMINATION PHOENIX

EQUATION NUMBER	RACE OR SUBMARKET VARIABLE INCLUDED IN EQUATION	SEMILOG EQUATION		LINEAR EQUATION	
		COEFFICIENT	t-STATISTIC	COEFFICIENT	t-STATISTIC
1	Spanish American household	-0.018+	1.602	-2.417+	1.637
	Black household	-0.069**	3.768	-7.269**	3.104
2	Proportion of households in Census tract that are Spanish American	-0.068	1.543	-10.270+	1.775
	Proportion of households in Census tract that are black	-0.152**	3.899	-18.521**	3.648
3	Spanish American submarket (Census tract > 50% Spanish American)	-0.006	0.251	-0.929	0.329
	Mixed submarket (Census tracts 20 to 50% Spanish American)	0.019	1.420	1.510	0.872
4	Spanish American submarket in central city	-0.025	0.791	-3.159	0.758
	Spanish American submarket in suburbs	0.009	0.293	-0.054	0
	Mixed submarket in central city	0.006	0.247	-0.194	0.063
	Mixed submarket in suburbs	0.032	1.390	3.443	1.176
	White submarket in central city	-0.006	0.298	-0.605	0.245
5	Spanish American household in Spanish American submarket	0.022	1.072	1.277	0.482
	White household in Spanish American submarket	0.022	0.775	3.592	0.987
	Spanish American household in mixed submarket	0.013	0.731	-0.125	0.055
	White household in mixed submarket	0.052**	3.916	5.474**	3.199
	Spanish American household in white submarket	0.005	0.295	1.836	0.831

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a C* neighborhood with fewer than five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

NOTE The equations used to test for discrimination are exactly like those used to test for racial submarkets in Chapter 4. The derived factor scores (for Census tracts) do not include the variables percentage black or percentage Spanish American in the Census tract (see Appendix VII). The remaining independent variables are those listed in the final equations in Chapter 3.

* Significant at the 0.05 level

** Significant at the 0.01 level.

\+ Significant at the 0.10 level.

linear additive form. (See Appendix II for the basic equations.) Since the sample of black households is relatively small, it is not possible to define separate submarket variables for blacks, other than the proportion of black households in the Census tract.

Equation (1) includes dummy variables that represent Spanish American and black households. The coefficients of these variables suggest that both minority groups, but particularly black households, pay less than white households for comparable housing.

In order to assess whether price differentials exist across submarkets, Equation (2) includes the percentage of Spanish American and the percentage of black households in Census tracts, and Equation (3) includes dummy variables for predominantly Spanish American submarkets and mixed (Spanish American and white) submarkets. From these equations it does not appear that housing in a completely Spanish American neighborhood is less expensive than comparable housing in predominantly white neighborhoods. The variable for percentage of Spanish American households in the Census tract is significant in only the linear from of Equation (2). The dummy variable for the Spanish American submarket in Equation (3) is not significant. However, for black neighborhoods a very large price differential appears: according to the coefficient of "percent black" in Equation (2), housing in all black tracts is 15 percent (semilog equation) less expensive than in all white tracts.

As in Pittsburgh, the variables in Equation (4) for Phoenix have been specified to include proxy variables for potential transition (mixed) submarkets as well as to represent minority and white submarkets. (Again, submarkets are defined according to the proportion of Spanish American households in Census tracts.) Predominantly Spanish American neighborhoods exist both in the central city and in the suburbs, so that the mixed submarkets in both areas represent potential transition areas. (The excluded category in Equation (4) is white submarkets in suburban areas.) None of these variables is significant, however; the coefficient for the variable representing the mixed market in the suburbs exceeds its standard error, but it seems difficult to label this as a target area based on this slim evidence.

Finally, Equation (5) distinguishes both ethnicity of household and type of submarket in order to test for "pure" price discrimination--that is, where minorities pay more than whites for comparable housing in the same neighborhood. The only significant variable is the one that represents white households living in mixed markets; they appear to pay more than white households living in white submarkets.

The results of these tests for price differentials in Phoenix are somewhat difficult to interpret. Clearly, they do not support the hypotheses of the market separation model, because there is no evidence of price markups in minority submarkets. Neither are the results entirely compatible with the Bailey-Muth equilibrium models, because the evidence that white households pay a premium to maintain segregation is mixed. Relative to Spanish American and black households, white households generally appear to pay somewhat more for comparable housing. According to the equilibrium model, however, whites in boundary (racially mixed) areas should pay less than white households inside the white submarket; Equation (5) indicates the opposite condition in Phoenix.

Again, the problem may be that the definition of submarkets on the basis of the racial distributions of Census tracts is too imprecise to adequately reflect "different" neighborhoods. The equilibrium models appear to assume that white and minority submarkets are almost completely segregated and that white households will occupy common boundary areas only when offered a discount relative to the white submarket. Neither condition holds in Phoenix. Given the present definition of submarkets, it is not possible to determine whether white households in Phoenix pay a premium to maintain distance from Spanish American or black households.

Another possible explanation is that certain omitted variables are systematically correlated with attributes of minority housing. If, for example, certain characteristics of lower-quality units were omitted, the coefficients for the variables that represent minority households in Equation (1) would be biased downward.

The next subsection makes additional tests for racial price differentials in Pittsburgh and Phoenix in order to substantiate the tentative conclusions reached here.

Subsample Tests for Price Discrimination

As already mentioned, most of the previous analyses of price discrimination are based on the common equation approach: an hedonic equation is estimated for an entire metropolitan area, and the coefficients of the race or submarket variables included in the equation are presumed to measure an "average" price differential that uniformly affects the prices of all types of dwelling units. This assumption is implicit in the market-wide equations estimated above in Pittsburgh and Phoenix. Such an approach may not adequately reflect the price effects of market segregation, however. Estimation of a single equation implicitly assumes that attribute prices are constant across the entire housing market; the analysis of market segmentation in Chapter 4 indicated that this is not in fact the case for some submarkets in Pittsburgh and Phoenix.

The effect of variation in attribute prices on minority housing prices is an important element of Kain and Quigley's analysis of segregation. The characteristics of ghetto housing may differ markedly from much of the stock in the white submarket. Although some types of dwelling unit quality or quantity may be obtained through modification of existing stock, conversion for other attributes is infeasible or prohibitively expensive. In addition, publicly or jointly produced attributes, such as public services and neighborhood amenities, are not easily alterable in the short run. All of these factors are likely to lead to different elasticities of supply for different attributes. For ghetto residents, higher quality dwelling unit or neighborhood attributes may bear a high premium or may not be available at all. Some types of housing bundles may therefore be more expensive inside the ghetto, and others may be cheaper.

In order to use the stratified equations to test for price discrimination, the estimated prices of equivalent housing bundles in different submarkets will be compared. These price estimates are obtained by solving the relevant subsample equations, using equivalent sets of attribute levels. Although numerous kinds of attribute bundles can be defined, an obvious choice is the "average" bundle consumed in each submarket.

Pittsburgh. The tests for price differentials indicated that a price premium exists for housing in the black submarket in Pittsburgh. To confirm this, the black submarket equation has been solved using the attribute levels for the average white submarket housing bundle. Similarly, the white submarket equation has been solved using the mean black submarket housing bundle. The existence of a price premium for ghetto housing is confirmed, but only for the average ghetto housing bundle. The estimated prices are given in Table 5-3. The results indicate that:

> If the average ghetto housing bundle were purchased in the white submarket, its cost would be about 4 or 5 percent less than in the ghetto submarket.
>
> If the average white submarket housing bundle were purchased in the ghetto submarket, its cost would be the same as in the white submarket.

It is important that these results hold for both linear and semilog forms. To the extent that the hedonic equation is misspecified, differences in estimated coefficients for different submarkets could simply reflect the misspecification; this is less likely to be the case when the results for both forms are so similar.

Analysis of the subsample equations for black and white households in Pittsburgh indicated that relative price differences existed between them. Table 5-4 shows the results of solving the equation for black households with the white household mean housing bundle, and vice-versa. No clear results are seen: black households may pay slightly less for the white bundles, but the difference is significant only in the linear equation.

Phoenix. A similar analysis has been done for Phoenix. Recall from Chapter 4 that evidence of market segmentation exists for submarkets in Phoenix. The analysis using full-sample equations indicates that Spanish American (and black) households appear to pay less than white households for housing.

Again, two sets of calculations have been made to assess these results. First, the equation for the Spanish American submarket has been solved using the mean attribute levels of the white submarket and the white submarket equation has been solved using the mean Spanish American market bundle. Second, the same calculations are made for the household groups and the average housing of the households.

Table 5-3

ESTIMATED PRICES FOR AVERAGE HOUSING BUNDLES
BY WHITE AND GHETTO SUBMARKET: PITTSBURGH

HOUSING BUNDLE	WHITE SUBMARKET PRICES	GHETTO SUBMARKET PRICES
	SEMILOG EQUATION (log rent, standard error, and dollar value of logarithmic estimates)	
White submarket bundle	4.685 (0.300) ($108.30)	4.689[a] ($108.75)
Ghetto submarket bundle	4.572[a] ($ 96.73)	4.624 (0.272) ($101.85)
	LINEAR EQUATION	
White submarket bundle	$113.23 (33.61)	$113.60[a]
Ghetto submarket bundle	$101.37[a]	$105.43 (26.30)

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a C* neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Estimated price obtained by solving submarket equations (stratified by Census tracts) with the mean attribute levels of the other submarket.

Table 5-4

ESTIMATED PRICES FOR AVERAGE HOUSING BUNDLES
BY WHITE AND BLACK HOUSEHOLDS: PITTSBURGH

HOUSING BUNDLE	WHITE HOUSEHOLD PRICES	BLACK HOUSEHOLD PRICES
SEMILOG EQUATION (log rent, standard error, and dollar value of logarithmic estimates)		
White household bundle	4.686 (0.299) ($108.39)	4.664[a] ($106.06)
Black household bundle	4.586[a] ($ 98.10)	4.604 (0.259) ($ 99.85)
LINEAR EQUATION		
White household bundle	$113.28 (33.40)	$109.31[a]
Black household bundle	$103.28[a]	$103.23 (26.77)

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a C* neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Estimated price obtained by solving submarket equations (stratified by race) with the mean attribute levels of the other submarket.

The results, given in Tables 5-5 and 5-6, are difficult to interpret. Spanish American households pay about 5 percent less for their own bundle than white households would pay for the same bundle. This is at least consistent in direction with the previous results. However, as seen in Table 5-5, the white submarket bundle can be purchased only at a large premium in the Spanish American submarket. It is clear that caution must be exercised in making this type of comparison. The housing bundles being compared must exist in both submarkets--the predicted price of a "luxury" unit in the barrio will be invalid if this type of housing is not available. The spread in actual mean rents between the white and the Spanish American submarkets is extremely wide, namely, a difference of over $50.00 (as compared with a spread of about $8.00 for the black and white submarkets in Pittsburgh). Thus, little confidence can be placed in the estimated price markup, especially when other evidence suggests that Spanish American housing costs for actual available housing are equivalent to or less than white household costs.

Table 5-5

ESTIMATED PRICES FOR AVERAGE HOUSING BUNDLES BY WHITE AND SPANISH AMERICAN SUBMARKET: PHOENIX

HOUSING BUNDLE	WHITE SUBMARKET PRICES	SPANISH AMERICAN SUBMARKET PRICES
	SEMILOG EQUATION (log rent, standard error, and dollar value of logarithmic estimates)	
White submarket bundle	4.946 (0.332) ($140.63)	5.083[a] ($161.26)
Spanish American submarket bundle	4.507[a] ($ 90.65)	4.471 (0.302) ($ 87.43)
	LINEAR EQUATION	
White submarket bundle	$147.88 (44.03)	$154.61[a]
Spanish American submarket bundle	($ 92.86)[a]	$91.42 (27.64)

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a C* neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Estimated price obtained by solving submarket equations (stratified by ethnicity) with the mean attribute levels of the other submarket.

Table 5-6

ESTIMATED PRICES FOR AVERAGE HOUSING BUNDLES
BY WHITE AND SPANISH AMERICAN HOUSEHOLDS: PHOENIX

HOUSING BUNDLE	WHITE HOUSEHOLD PRICES	SPANISH AMERICAN HOUSEHOLD PRICES
	SEMILOG EQUATION (log rent, standard error, and dollar value of logarithmic estimates)	
White household bundle	4.896 (0.351) ($133.75)	4.900[a] ($134.29)
Spanish American household bundle	4.716[a] ($111.72)	4.670 (0.365) ($106.64)
	LINEAR EQUATION	
White household bundle	$141.50 (44.62)	$142.62[a]
Spanish American household bundle	$118.81[a]	$113.90 (42.06)

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a C* neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. Estimated price obtained by solving submarket equations (stratified by Census tracts) with the mean attribute levels of the other submarket.

REFERENCES

Abt Associates Inc., Working Paper on Early Findings, Cambridge, Mass., January 1975.

Bailey, Martin, "Effects of Race and Other Demographic Factors on the Value of Single-family Homes," Land Economics, vol. 42, May 1966, pp. 215-20.

Becker, Gary, The Economics of Discrimination, Chicago, University of Chicago Press, 1957.

Berry, Brian and Robert Bednary, "A Hedonic Model of Prices and Assessments for Single-family Homes," Land Economics, February 1975.

Daniels, Charles, "The Influence of Racial Segregation on Housing Prices," Journal of Urban Economics, vol. 2, 1975, pp. 105-22.

Gillingham, Robert, "Place-to-Place Rent Comparisons Using Hedonic Quality Adjustment Techniques," Research Discussion Paper No. 7, Washington, D.C., Bureau of Labor Statistics, March 1973.

Haugens, Robert and A. James Heins, "A Market Separation Theory of Rent Differentials," Quarterly Journal of Economics, vol. 83, November 1969, pp. 660-72.

Kain, John, "The Journey-to-Work as a Determinant of Residential Location," Papers and Proceedings of the Regional Science Association, vol. 9, 1962, pp. 137-61.

__________, "Housing Segregation, Negro Employment, and Metropolitan Decentralization," Quarterly Journal of Economics, vol. 62, June 1972, pp. 263-77.

Kain, John and John Quigley, Housing Markets and Racial Discrimination, New York, National Bureau of Economic Research, 1975.

King, Thomas and Peter Mieszkowski, "Racial Discrimination, Segregation, and the Price of Housing," Journal of Political Economy, vol. 81, May-June 1973, pp. 590-606.

Lapham, Victoria, "Do Blacks Pay More for Housing?" Journal of Political Economy, vol. 79, November-December 1971, pp. 1244-57.

Laurenti, Luigi, "Property Values and Race," Readings in Urban Economics, edited by Matthew Edel and Jerome Rothenberg, New York, MacMillan Co., 1972.

Lieberson, Stanley, Ethnic Patterns in American Cities, Glencoe, Illinois, Free Press of Glencoe, 1963.

McEntire, Davis, Residence and Race, Los Angeles and Berkeley, University of California Press, 1960.

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

Myrdal, Gunnar, The American Dilemma: The Negro Problem and Modern Democracy, vol. I, New York, Harper and Row, 1962.

Pascal, Anthony, "The Analysis of Residential Segregation," Financing the Metropolis, Edited by Crecine, Beverly Hills, California, Sage. Publications, 1970.

Quigley, John, "Racial Discrimination and the Housing Consumption of Black Households," Patterns of Racial Discrimination, Vol. I: Housing, edited by George M. vonFurstenberg, Lexington, Mass., D.C. Heath, 1974.

Rapkin, Chester, "Price Discrimination against Negroes in the Rental Housing Market," Essays in Urban Land Economics, Los Angeles and Berkeley, University of California Press, 1966.

Ridker, Ronald and John Henning, "Determinants of Residential Property Values with Special Reference to Air Pollution," Review of Economics and Statistics, May 1967.

Schnare, Ann, Externalities, Segregation, and Housing Prices, Washington, D.C., Urban Institute, 1974.

Stengel, Mitchell, Market Separation and Racial Price Differentials: The Ghetto Housing Market, unpublished, Lansing, Michigan State University, 1973.

Taueber, Karl and Alma Taueber, Negroes in Cities: Residential Segregation and Neighborhood Change. Studies in American Negro Life, New York, Atheneum, 1969.

APPENDIX I

THE INTERPRETATION OF HEDONIC INDICES AS MEASURES OF HOUSING QUALITY[1]

This appendix discusses the formal rationale for and limitations of hedonic indices as a measure of housing change. Hedonic indices are an important analytic tool for assessing changes in participant housing during the experiment. They are not the only tool. They complement rather than replace summary measures based on specific policy-relevant features or individual satisfaction. And, like any aggregate index, they cannot fully replace more detailed examination of individual dwelling unit features. Nevertheless, given the limitations of available indices based either on policy standards or on individual satisfaction, and the plethora of relevant housing features, hedonic indices are potentially the most useful single measure of housing now available.

The basic difficulty in measuring housing is that it involves a collection of many different attributes. Individual features can and should be considered, but the large number of features and the many alternative ways of describing them require some summary measures as well. Summary measures can be constructed from at least two different viewpoints--that of social policy and that of individual well-being. Social policy ratings attempt to evaluate housing in terms of externally set requirements. These requirements are usually based on notions of basic amenities, such as indoor plumbing and features necessary for safety and health, or on presumed externalities produced by decent housing, such as improved appearance, reduced crime and disease, and so on. Measures based on individual satisfaction, on the other hand, are basically concerned with the extent to which an individual household's housing needs are met. At their most ambitious, individually motivated measures attempt to abstract from particular households and to identify a common scale of housing needs and adequacy that reflects a general consensus about what constitutes "good" housing. As discussed below, hedonic indices may be seen as a special instance of this latter approach.

---

[1] This appendix was written by Stephen D. Kennedy and Sally R. Merrill.

The problem in developing ratings based on social policy considerations is lack of consensus. The Minimum Standards used in the Demand Experiment, for example, are developed from the American Public Health Association's Recommended Housing Maintenance and Occupancy Ordinance (as revised in 1971) and are similar to standards used in other housing programs, such as Section 8. Nevertheless, there is little question that individual policy makers would quarrel with the omission or inclusion of specific standards, or with the relative weight given to, for example, floor condition, safe electrical wiring, or presence of adequate plumbing.

On the other hand, measures based on individual satisfaction with dwelling unit or neighborhood may lack credibility. Consider, for example, an individual's expressed satisfaction with his or her neighborhood. The measure itself is subject to a variety of limitations (such as the common observation that people tend to ratify their present situation, and especially their recent choices, by claiming satisfaction). More important, the subjective nature of individual satisfaction may be unpersuasive on at least two grounds. First, individuals may be dissatisfied with their housing not because it is inadequate, but simply because it is unsuited to their unique needs (most obviously, a dwelling unit that is too large or too far from a new job). Second, an individual's dissatisfaction with housing may be suspected of reflecting unreasonable desires.

One approach to these problems is to attempt to build a measure of housing by identifying an underlying structure of housing tastes or needs common to all individuals. Such approaches are epitomized by latent trait models and their associated factor analytic approaches. The problems with this approach are twofold. First, no observable variable validates the derived structure: Because the identification of traits is dependent on prior restrictions, it is difficult to prove that the factors do indeed identify some common structure. This problem can be substantially overcome in cases where the identified factors possess strong surface plausibility or are replicated in different situations. Second, and more fundamentally, the latent traits, even if identified, are difficult to interpret. Once housing has been reduced to, for example, seven different dimensions, there is still no accepted scale for the units and no immediate way to understand the importance of a unit change. Justification and interpretation must ultimately rest on the experience built up by repeated applications of the factors to various outcomes, which establish both their significance in determining outcomes of interest and the magnitude of differences in outcomes associated with differences in factors. This sort of justification requires substantial time to develop, however.

If there is some observed variable that is commonly thought to be correlated with housing adequacy, it may be used to interpret the derived latent traits. Alternatively, housing attributes may be related to it directly, without attempting to identify an underlying structure. Indeed, this constitutes one approach to the interpretation of hedonic indices: Based on the assumption that people will generally pay more for a dwelling only if it is better, different attributes are weighted according to the way in which they affect the market value of the unit. The total value of the unit's attributes is then its estimated normal market, or hedonic, value. This value is different from the unit's actual rent, which may reflect a variety of nonhousing factors, including the effects of inflation over time and the careful shopping or luck of individual households in finding especially good deals.[1]

In fact, the conditions under which hedonic indices can be interpreted in this way are stringent and probably not met. Hedonic indices of housing cannot reasonably be claimed to identify either a common set of consumer preferences and housing needs, or the underlying housing supply costs for different sorts of housing. Under certain circumstances, however, hedonic indices can be thought of as identifying common agreement not about whether one house is better than another, but rather about whether it is worth more and in some sense provides "more" housing.

The idea of "more" or "less" housing is best represented by the common habit of referring to a $40,000 house or a $200 apartment (or, for automobiles, to a high, medium, or low-priced car). This in effect characterizes houses (or cars) in terms of their normal market cost. A particular $40,000 house may sell for more or less than $40,000, and it

[1]In addition, estimated hedonic values will of course differ from actual hedonic values due to errors in estimation.

may be more or less suited to a particular household's needs than another house. But there is, in conversation, the idea that it is "more" house than a $20,000 house and in some very loose sense, a better house. Put another way, if an individual with a $10,000 house were to purchase a $20,000 house, he would seek to purchase a "better" (for him) house. Hedonic indices provide a more-detailed and objective version of this sort of characterization of housing, but their strengths and weaknesses can still be understood in terms of the strengths and weaknesses of such characterizations.

Most obviously, a $200 apartment may mean quite different things in New York City as compared with Springfield, Massachusetts, or in 1977 as compared with 1937. In this case, an hedonic index of housing will also mean different things. As discussed below, this problem can be partly overcome by using the weights from one market to index housing services in other markets. If the true weights in the two markets differ only proportionally, this comparison will yield an exact index. Otherwise, one market's weights will generally overestimate the level of housing services in other markets relative to the first, and will misestimate changes or differences in housing services within the other market in an unknown way.[1] It is therefore important to examine the extent to which hedonic weights in two situations differ by more than a factor of proportionality.

Beyond this, the theory of hedonic indices assumes that there is an adequate description of dwelling units and of the relationship between a unit's features and its normal market rent. In terms of actual descriptors, theory provides little guidance, beyond the rule that variables should reflect unit attributes or nonhousing factors (such as being related to

---

[1] Using the attribute weights from market A to index housing in market B in effect estimates what a given market B unit would cost in market A. If costs in the two markets differ only by a factor, this simply rescales the market B index; the ordering of units remains the same using either market A or market B weights. If, however, the weights differ in a nonproportional way, they may give different orderings of units within each market. Furthermore, between the two markets, market A weights will tend to overvalue market B housing, since, for a given real level of housing services, a unit in market B will tend to have more of the attributes that are inexpensive in market B as compared with market A.

the landlord) that might affect a unit's actual rent. In terms of the appropriate form of the relationship between attributes and normal market rent, theory at least suggests that this will often be nonlinear, except in cases where the attributes are themselves individually bought and sold (see Lucas, 1975, and Section I.2 below).

These issues are discussed further in Section I.1, which summarizes the general theory of hedonic indices. Section I.2 then reviews some of the recent literature on hedonic indices, much of which has been aimed at removing misconceptions about what hedonic indices do and do not measure.

The remainder of this appendix is very much rooted in the economic theory of consumer choice. While the presentation attempts to develop a more general intuitive content, the concepts involved are motivated by attempts to define the concept of commodities within economic models. Without prior exposure to such models, the arguments may be hard to follow; common English terms are used with special meanings, and the logic and motivation of the discussion rest on these models. Since hedonic indices are largely developed and used by economists, however, it is worthwhile to present their use and limitations within the framework of economics.

I.1 THE USE OF HEDONIC INDICES TO MEASURE HOUSING

The hedonic approach to the analysis of consumption rests on the assumption that goods may be disaggregated into sets of basic characteristics, and that the characteristics of the particular good, rather than the good itself, constitute the arguments of the consumer's utility function. As summarized by Zvi Griliches, the parametric version of the hedonic technique asserts the existence of a "reasonably well-fitting" relation between the price of the composite commodity (or good) and the levels of its characteristics (Griliches, 1971). The coefficients that result from estimation of this relationship, usually a regression, are referred to as the shadow, or implicit, prices of the characteristics. Traditionally, the hedonic methodology has been used in the construction of price indices, a use developed by Andrew Court (1939). The technique was revived later by Griliches and others and has been applied in many interesting ways.[1]

[1] Griliches (1971) provides a collection of studies that describe the applications and issues of hedonic research.

Hedonic indices have a strong surface plausibility. Housing is not a single, well-defined commodity, but rather a collection of attributes such as number of rooms, structural soundness, adequate plumbing, pleasing design, neighborhood services and amenities, and indeed neighbors themselves. Hedonic indices in effect assume that, in the market as a whole, rent is strongly related to the quantity and quality of housing services that a unit offers. Thus, by regressing rent on attributes, one can derive a set of weights which reflect each attribute's contribution to the value of the unit. In particular, hedonic regressions can be used to sort out the value of housing services from other factors that influence rent, such as individual shopping ability, conditions of tenure, or racial discrimination.

On the other hand, the adequacy and quality of a house appear to lie very much in the eye of the beholder. Different people rate units very differently, depending on their needs and background. The relative importance of air-conditioning and central heating, for example, are very different in Pittsburgh and Phoenix. How then can housing be considered a well-defined entity that can be measured by a single index?

One approach to the definition of housing quality is to assume some underlying psychological agreement about what constitutes good or bad housing. Say that each individual possesses a preference ordering, $U(x,z)$, defined over some set of housing attributes, x, and other goods, z. Then the index of housing attributes for each individual is written as

(1) $$h(x;\bar{z}) = U(x,\bar{z}),$$

where $h(x;\bar{z})$ = the index of housing services, given other consumption

x = the vector of housing attributes

$\bar{z}$ = some fixed vector of other consumption.

The function $h(x;\bar{z})$ indexes housing in the sense that if $h(x_1;\bar{z})$ is greater than $h(x_o;\bar{z})$, then x_1 is preferred to x_o (given other consumption, $\bar{z}$).[1]

[1] The usual assumption is that h and U only order different housing bundles--that is, h and U are only defined up to a monotonic transformation, since any monotonic transformation will leave the ordering of bundles unchanged. Thus the index, h, only says whether one house is better than another, not how much better.

An index may be constructed independent of z if the individuals' ranking of alternative values of x in U(x,z) is independent of z. Formally, this can be written as

$$(2) \qquad U(x,z) = U(g(x),z).$$

The utility function of Equation (2) is called "separable" in terms of housing and nonhousing goods. If U is increasing in g, then x_1 will be preferred to x_0 if $g(x_1) > g(x_0)$, regardless of the value of z.

The fact that people readily talk about housing apart from other goods at least suggests the reasonableness of separating housing and nonhousing preferences. For the variables used in this report, however, separability would require, for example, that a household's relative preference for an extra bedroom or a garage is unaffected by whether or not it owns a car.[1]

Even if preference orderings are separable, the index, h(x), will not be the same for all individuals unless individual tastes are all the same. Again, this may be a reasonable assumption (at least in application) if attributes are appropriately defined. Thus, for example, it may be reasonable to suppose that there is some quality "interior climate" over which preferences are identical. The actual variables available, however, are the presence and adequacy of central heating and air-conditioning. Relative preferences for these two forms of climate control can hardly be the same in Pittsburgh and Phoenix. Likewise, there may be some common underlying quality of roominess. Tastes for square feet or number of rooms, however, would be expected to vary with household size and composition (as well as exterior climate).[2]

If the attempt to identify a common underlying psychological agreement by using available descriptors is problematic, another approach is to

---

[1]A necessary and sufficient condition for the separability defined by Equation (2), due to Leontieff, is that the marginal rate of substitution between any two elements of x (the ratio of marginal utilities) be independent of the elements of z.

[2]These sorts of problems indicate why the validation of latent trait models, discussed at the beginning of this appendix, is so important and so difficult. Given available descriptors, the definition of latent traits for housing would be expected to shift from city to city by demographic characteristics. This means that factor scores, for example, should change across demographic groups and cities even though the factors may seem, on interpretation, to be identifying the same element. Absent internal evidence of consistency, it becomes even more important to develop external tests.

restrict the available choice set so that differences in individual tastes are effectively suppressed. This is the basis of a remarkable theorem due to Hicks, the Composite Commodity Theorem. Say that the relative prices of some subset, A, of goods are fixed--that is, the price of each good in the subset rises or falls proportionally; then, under the conditions of utility maximization, every individual will act as if the subset of goods were a single composite commodity, α, defined by:

$$\alpha = \sum_{i \in A} \left(\frac{P_i}{P_\alpha}\right) X_i \tag{3}$$

$$P_\alpha = \sum_{i \in A} P_i . \tag{4}$$

As long as the subset of prices rises or falls proportionally, the weights that define α (the P_i/P_α) remain fixed. Thus α provides an index of the subset $(X_1 \ldots X_r)$, and P_α provides an index of the subset prices.

It is important to understand what this theorem does and does not say. It does not define a single physical commodity that all individuals will purchase. The composition of the composite commodity in terms of the amounts of the individual goods involved (the X_i) may vary among individuals and, for any single individual, as income or price levels change. The theorem does maintain that in considering behavior we need not define any ultimate commodities: people can be thought of as deciding the level of α and then, behind the scenes as it were, allocating α among its individual elements.

Put another way, the composite commodity measures the quantity of food or housing an individual buys, not its quality. For example, if individual A buys two bags of groceries, one for $5 and one for $10, individual B may prefer the beer and pretzels that made up the first bag to the soybeans, spinach, and cabbage that made up the second. But in a general sense it would be agreed that the second bag contains more groceries. It has a higher value in the sense that if individual B were to buy $10 worth of groceries, he would get more (or better) groceries--for him--than if he bought only $5 worth. The Composite Commodity Theorem in effect provides a rigorous basis for the notion of talking about a $25 bag of groceries or a $40,000 house; it says that $25 worth of groceries does in fact refer to the cost of a composite good called "groceries" and does indeed measure the amount of "groceries" up to a scale factor (the price).[1]

Hedonic indices involve a further step: goods are seen as bundles of attributes. Thus, the houses in a particular city are seen not as hundreds of thousands of unique commodities, but rather as different combinations of a limited set of attributes. The Composite Commodity Theorem can be applied to the underlying attributes as well as to individually marketed commodities. If the relative prices of a subset of attributes are fixed, then the attributes may be formed into a composite attribute bundle. There is, however, no reason to assume that attributes will have prices in the usual sense. Attributes are embodied in marketed goods, so that the cost of an attribute set, x, is given by:

(5) $$C(x) = \min p_t' t \text{ s.t. } F(t) \geq x,$$

where x = the vector of attributes

t = the vector of marketed commodities

p_t = the vector of market prices

F = the function that maps t into x.

The market cost function for the attributes, C(x), will be linear only under very special conditions. Most obviously, if each marketed good contains given amounts of attributes per unit, and if there are the same number of marketed goods as attributes, then

(6) $$x = Qt$$

$$p_x = p_t' Q^{-1}$$

$$C(x) = p_x' X,$$

where $\left\{Q_{ij}\right\}$ = the amount of the i^{th} attribute contained in a unit of the j^{th} marketed commodity (assumed to be nonsingular).

---

[1] The application of the Composite Commodity Theorem to hedonic indices of housing services is one example of a much larger problem. There is an abundance of commodities; there are dozens of brands of soap or models of cars or types of houses. Further, each car or house, at least, is potentially unique. Yet we are accustomed to think in terms of broad categories such as cars, housing, or even simply income. For economists, at least, this is not simply verbal sloppiness. Nor does it require assumptions about regularity of tastes. It can simply reflect the underlying unity of categories of goods engendered by a unity of changes in price.

But this is a trivial case, since the point of considering attributes was to reduce dimensionality. Indeed, to the extent that there are more varieties of goods than attributes, this suggests that individuals are not efficient producers of attributes, that it pays to have firms produce different bundles. Thus, as Lucas (1975) points out, if the Q-matrix in Equation (6) is singular (that is, if there are more commodities than attributes), then the cost function, C(x), will be nonlinear (specifically a polygonal arc concave to the origin), except in the degenerate case in which some subset of commodities dominates (that is, in which there is no reason for there to be any more commodities marketed than attributes). In addition, Rosen (1974) points out that the formulation of Equation (6) is itself too simplistic; for example, two six-foot cars cannot be combined to give a twelve-foot car.

Fortunately, the Composite Commodity Theorem does not depend on linear cost functions. A composite commodity, h(x), can be constructed as long as the cost of purchasing a set of attributes, x, can be expressed as:

(7) $$C(x) = \theta f(x) g(z),$$

where θ = a shift parameter

$g(z)$ = some function (possibly constant) of the other goods

$f(x)$ = a fixed function of the attributes.[1]

---

[1]This can be proved as follows. Consider any nondecreasing index, h(x). Define

(a) $$W(\alpha,z) = \max_{\{x\}} U(x,z) \text{ s.t. } h(x) = \alpha.$$

This defines a preference ordering over (α,z) and a set of correspondences between $\hat{x}$, the solution to Equation (a), and (α,z). If h(x) is not convex, it may coincide with the indifference curves of U(x,z) at multiple points. If this is the case, a function $\hat{x}(\alpha,z)$ may be defined by choosing the least cost value among the x solutions:

(b) $$\max_{\{\alpha,z\}} W(\alpha,z) \text{ s.t. } D(\alpha,z) = Y,$$

where $D(\alpha,z)$ is defined by

(c) $$D(\alpha,z) = E(\hat{x}(\alpha,z),z),$$

where E(x,z) is the cost function for purchases (x,z). The index, h(x), can be considered a composite commodity if the solution to Equations (a) and (b) yields the same solution for (x,z) as

(footnote continued on next page)

(footnote continued from previous page)

(d) $$\max_{\{x,z\}} U(x,z) \text{ s.t. } E(x,z) = Y.$$

By the Envelope Theorem and the first order conditions for Equation (a),

(e) $$\frac{\partial W}{\partial \alpha} = \mu = \left(\frac{\partial U}{\partial x_i}\right)\left(\frac{\partial h}{\partial x_i}\right)^{-1} \; ; \; \frac{\partial W}{\partial z_i} = \frac{\partial U}{\partial z_i} .$$

Substituting Equation (e) into the first order conditions for Equation (b) gives

(f) $$\frac{\partial U}{\partial x_i} = \eta\left(\frac{\partial h}{\partial x_i}\right)\frac{\partial D}{\partial \alpha} \; ; \; \frac{\partial U}{\partial z_i} = \eta\frac{\partial D}{\partial z_i} \; ; \; D = Y,$$

whereas the first order conditions for Equation (d) are

(g) $$\frac{\partial U}{\partial x_i} = \eta\frac{\partial E}{\partial x_i} \; ; \; \frac{\partial U}{\partial z_i} = \lambda\frac{\partial E}{\partial z_i} \; ; \; E = y.$$

Assume that the cost function, E, can be written

(h) $$E(x,z) = \theta f(x)g(z) + k(z),$$

and define the composite commodity index, h(x) by

(i) $$\alpha = h(x) = \frac{f(x)}{f(1)} ,$$

and the cost function D by

(j) $$D(\alpha,z) = p_\alpha \alpha + k(z),$$

where

(k) $$p_\alpha = \theta f(1)g(z).$$

Then Equations (f) and (g) can be rewritten

(f)′ $$\frac{\partial U}{\partial x_i} = \eta\theta g(3)\frac{\partial f}{\partial x_i} = \frac{\partial U}{\partial z_i} = \eta f(x)\frac{\partial g}{\partial z_i} + \frac{\partial k}{\partial z_i} ; \; \theta f(x)g(z) + k(z) = y$$

(g)′ $$\frac{\partial U}{\partial x_i} = \lambda\theta g(3)\frac{\partial f}{\partial x_i} \; ; \; \frac{\partial U}{\partial z_i} = \lambda f(x)\frac{\partial g}{\partial z_i} + \frac{\partial k}{\partial z_i} \; ; \; \theta f(x)g(z) + k(z) = y,$$

which are identical. Thus Equation (h) is sufficient. On the other hand, Equations (e) and (f) require that

(l) $$\frac{\partial h}{\partial x_i} = \left(\frac{\lambda}{\eta}\right)\left(\frac{\partial D}{\partial \alpha}\right)^{-1}\frac{\partial E}{\partial x_i} .$$

Since h must be independent of z and since, because tastes are unrestricted, Equation (l) must hold for all values of x and z, the Equation (h) must also be necessary. Thus the basic requirement for indexing x across individuals is that all individuals face the same function of the "separable" form given by Equation (h).

The form of Equation (7) allows housing costs to depend on nonhousing consumption, z, as well as on housing consumption. In practice, hedonic indices for housing are usually estimated without considering nonhousing consumption. Thus the empirically appropriate form for Equation (7) is

$$C(x) = \theta f(x). \tag{8}$$

Equation (8) simply requires that the cost of a given unit not change as other consumption (such as food purchases) changes. This requirement may seem innocuous at first glance, but is in fact important. Most obviously, Equation (8) requires that the attributes, x, not be produced by the omitted goods, z. This is in effect a technical, or market separability, condition, which serves the same function as the separability condition on preferences in Equation (2). The condition is stronger than a simple separability of attributes, however. Many urban economists would argue, for example, that the price of housing and indeed the relative price of various attributes changes with distance from the workplace and shopping centers. But this means that C(x) must be written as

$$C(x) = \theta f(x,t), \tag{9}$$

where t represents the location of the unit. The hedonic index for housing cannot be separated from location.[1]

The estimation of hedonic indices in effect attempts to estimate the weights for the composite commodity of quality attributes. Of course, if rent were determined only by housing quality, it could be used as a direct measure of the composite housing bundle. Hedonic estimation is used to sort out the market value of quality attributes from the effects

---

[1] It may be useful to distinguish two different problems here. If there is a price gradient along which relative prices shift, then that gradient must be included in estimating the hedonic index. This is a market cost descriptor. In addition, however, the travel costs associated with a particular location will vary from individual to individual, depending on exact work location, shopping needs, type of transport, and so forth. As long as an individual can purchase a given amount of "travel cost" for any housing bundle, "travel cost" can be regarded as another commodity (part of z) and will enter the housing cost equation as g(z) in Equation (7). In this case, the hedonic index is preserved. This preservation requires, in the extreme, that every housing bundle be available at every location (or, more exactly, that every relevant bundle be available at any given travel time from relevant work and shopping centers).

of individual shopping behavior, tenure conditions, and other nonquality factors, as well as the effects of price changes over time.[1]

In addition, hedonic indices can be used to compare housing in different markets with different housing price structures. The composite rationale depends critically on the assumption that the relative attribute weights in the hedonic regression are fixed. Yet these weights will differ over time, between cities, and across submarkets within cities if the attribute cost function differs. If attribute costs only differ proportionally, then the composite commodity is of course maintained. The original weights can be used in both situations. This in effect simply adjusts for differences in the price level between the two times, cities, or submarkets. If the relative weights change, the composite commodity changes as well and can no longer be directly compared with the original composite; the two are not totally unrelated, however.

The problem of comparing housing composites across different markets with different attribute weights is essentially the same as the problem of constructing price indices. A price index is simply a deflator that attempts to scale the overall composite commodity so that it is comparable to income under some set of base prices. The properties of such indices are well known and apply directly to comparison of housing bundles.

Assume that nonhousing goods are fixed. The individual then selects housing attributes to maximize $U(x,\overline{z})$--the conditional housing index of Equation (1)--subject to the constraint that housing expenditures, $C(x)$, must be less than income net of nonhousing expenditures, $\hat{y}$.[2] Under reasonably general conditions this defines the housing index as a function of costs and expenditures by

---

[1]As discussed in Chapter 2, the resulting index for any given time period may be formed either by using the estimated hedonic function, excluding nonquality factors, or by using the estimated coefficients for nonquality factors to adjust rent. The choice of approach depends on whether the residual from the hedonic estimates is thought of as omitted housing attributes or as omitted nonquality factors. Indexing over time would require that adjusted rent be further adjusted for inflation.

[2]An unconditional index can be defined if the preference ordering is separable, as in Equation (2). For further discussion in terms of price indices, see Pollack (1976).

$$(10) \qquad \emptyset(C,\hat{y}) = \max_{\{x\}} U(x,\bar{z}) \text{ s.t. } C(x) \leq \hat{y},$$

where y is income net of nonhousing expenditures. The indirect utility function, $\emptyset$, is homogeneous of degree zero (since proportional shifts in C and y leave the feasible set of x unchanged), so that $\emptyset$ can be written as $\emptyset[\frac{C(x)}{C(1)} ; \frac{\hat{y}}{C(1)}]$. But $\hat{y}/C(1)$ is simply the composite commodity that indexes $U(x;\bar{z})$, given the cost structure C.

Alternatively, given any cost structure, C_1, and expenditure, $\hat{y}_1$, an index, I, of $U(x,\bar{z})$ can be defined relative to some other cost structure, C_0, by

$$(11) \qquad \emptyset(C_0,I) = \emptyset(C_1,y_1).$$

If the utility function, $U(x,z)$, were known, then the index, I, could be derived. If housing expenditures alone are known under the two-cost situation, it is clear that

$$(12) \qquad C_0(x(C_1,y_1)) > I,$$

since with net income $C_0(x_1)$ the household could at least purchase x_1 and generally would not (that is, the household would achieve a higher level of utility by purchasing some other point in the new constraint set defined by C_0).

Thus, if costs change over time from C_0 to C_1, the use of the base period weights to value housing at t_1 will tend to overvalue the housing at t_1, in terms of t_0 costs. On the other hand, between the two cities, Phoenix weights will overestimate the value of Pittsburgh bundles relative to Phoenix bundles, and Pittsburgh weights will overestimate the value of Phoenix bundles relative to Pittsburgh bundles. Note that nothing is said about the relation between real housing change in Phoenix and the change in an index that uses Pittsburgh weights. The Pittsburgh-based index will overvalue both the base-period and second-period index in Phoenix.[1] Thus, there is no apparent way to bound the absolute change in

---

[1]Economists will recognize these as the standard results of price index theory, except that the results posed for hedonic indices refer to the nonlinear composite commodity, $C(x)/C(1)$. That is, values are taken from the cost function and not from marginal prices. It is apparent that the revealed preference derivation of these propositions applies to nonlinear as well as to linear cost functions. See Samuelson (1947), pp. 141-163.

Phoenix by using Pittsburgh's weights. However, since the hedonic index has an arbitrary scale (dollars), this may not be too great a loss. Percentage changes using each city's own weights are directly comparable.

The same problem applies to submarkets within a city. An index based on one submarket will overvalue the other. Further, changes in an index based on one submarket have an unknown relation to changes in the true index for other submarkets.

It is worth noting that submarkets may pose no problem at all in some cases. If, for example, the cost structure of attributes is quite different for multifamily and single family units, such interactions may be incorporated directly into the estimated hedonic regression. As long as individuals have equal access to both types of units, the overall composite commodity can include the choice of building type and the resultant shift in attribute weights within that type. A problem arises only when households do not have equal access to different submarkets. Thus submarkets maintained by racial discrimination or by different concentrations of employment opportunities for different types of households can in theory pose serious problems.[1]

Nothing can be done about these problems without attempting to develop estimates of the utility function, except to point out that the index is only approximate under these circumstances, and then to assess the extent to which nonproportional differences actually exist in the cost structure over time, between cities, or among submarkets within cities.

---

[1] Limited access may not simply reflect barriers to specific locations. It may also reflect the absence of certain housing combinations. The composite commodity rationale implicitly assumes a certain continuity in the available combinations of attributes. But this may not apply to housing markets. Thus, the good schools of the suburbs may be unavailable to a poor household because they only come packaged with large houses and two-acre lots. Such discontinuities, if they exist, could force people to corner solutions (that is, to situations where the shadow costs are different from the shadow costs in the available space). In this case, the hedonic composite commodity is reduced to a price index (since the "real" shadow costs are shifting as individuals move from internal to corner solutions).

I.2 FURTHER CONSIDERATIONS IN THE USE OF HEDONIC INDICES

The previous section presented a rationale for hedonic indices in terms of the notion of composite commodities. Hedonic regressions were essentially regarded as estimating the current market cost of bundles of housing attributes; standard price index theory was then used to develop measures of real housing from this market cost function.[1] This section reviews some recent literature that has helped remove various misconceptions about what hedonic cost functions do and, in particular, do not mean.

Many empirical analyses that use the hedonic technique give as theoretical support either Lancaster's (1966) or Muth's (1969) model of household production. The essence of this approach has been summarized by Lancaster (1966) as follows:

> The good [commodity], per se, does not give utility to the consumer; it possesses characteristics or attributes, and these attributes give rise to utility.
>
> In general, a good possesses more than one attribute, and many attributes are shared by more than one good.
>
> Goods in combination may possess attributes different from those pertaining to goods separately.

In Lancaster's model, consumers maximize utility in terms of attributes, $U(x)$, subject to the cost constraint, $C(x) \leq y$. The cost function, $C(x)$, is defined by

$$C(x) = \underset{\{t\}}{\mathrm{Min}}\ p_t' t \text{ s.t. } x \geq Qt, \tag{13}$$

where
x = a vector of attributes
t = a vector of commodities
p_t = a vector of commodity prices
Q_{ij} = the amount of the i^{th} attribute in the j^{th} commodity.

Lancaster's major concern has been to analyze consumer decisions and to

---

[1] This seems to be in the spirit of Adelman and Griliches (1971), Griliches (1971), and Triplett (1971), though details differ.

assess the degree of complementarity or substitutability among goods on the basis of their shared attributes.[1]

Much of the recent literature has been devoted to developing a better understanding of the attribute-commodity model and its implications for hedonic indices. Thus, for example, Lucas (1975), Muellbauer (1974), and Murray (forthcoming) explore conditions on preferences and attribute-commodity mappings that allow estimation of subindices for commodity groups such as housing and food. Likewise, Lucas (1975) and Rosen (1974) begin to develop insight about why there are more commodities than attributes and what this suggests about the attribute cost function.

In terms of hedonic indices, the discussions in the literature seem to have centered on three points:

> First, hedonic cost functions generally do not identify either underlying tastes or supply functions, any more than do current commodity prices (Lucas, Rosen, Kain, and Quigley).
>
> Second, hedonic price indices are no less subject to aggregation problems than are commodity price indices (Lucas, Muellbauer).
>
> Third, nonlinearity of hedonic cost functions presents some special problems, particularly when marginal costs are treated as if they were market prices (Lucas, Muellbauer, Murray).

Each of these points is discussed briefly below, first, in the context of the classical budget constraint (a linear function of market prices), and second, with reference to the special features of hedonic indices.[2]

---

[1] In an earlier article, Irma Adelman and Zvi Griliches also described the theoretical foundations of the hedonic approach in terms of a household production model, along the lines developed by Lancaster (Adelman and Griliches, 1961).

[2] It is not completely clear why these points were ever in dispute. They seem to be worth recapitulating, however, since the literature on hedonic indices usually starts in medias res and does not draw out the connection between hedonic indices and other price indices.

In classical theory, consumers maximize a utility function, $U(x)$, subject to a budget constraint, $p'x \leq y$. Consumers range themselves across the budget constraint due to variations in taste or income, and the observation that each consumer has an indifference curve tangent to the budget line at his or her consumption point provides only limited information about the structure of the entire indifference surface.

As one would expect, the same is true of hedonic cost functions (Rosen, 1974). In particular, although nonlinear cost functions provide a variety of marginal costs (tangent lines), this does not provide any more information than the fixed-marginal-cost linear budget constraint about consumers' indifference surfaces. Observations do not represent a set of alternative outcomes under different prices, since consumers face the cost function, rather than the planes tangent to it, as a constraint. Thus, for example, if individual tastes are all the same, individuals with different incomes may face different marginal prices, but observations of consumption under these different marginal prices give no more information about tastes than do observations of consumption under fixed marginal prices.[1,2]

Likewise, observation of current market prices gives information only on the tangent line to the production surface at one point, and then only under appropriate equilibrium conditions. Similarly, observation of attribute costs gives an envelope of tangents for the given overall level of attribute production (Rosen, Lucas). Thus, although the attribute cost function provides information on the market cost of a variety of attribute bundles for a given level of overall attribute production,

---

[1] As with the linear budget constraint, under appropriate assumptions about tastes, variations in consumption patterns across income groups provide information on the expansion path for attribute consumption. Indeed, given appropriate concavity (from below) for the iso-cost functions, these demand properties may be analyzed in terms of income and the "prices" given by the tangent plane at the point of consumption.

[2] The really interesting problem is reflected in variations of consumption among individuals with the same income, reflecting variations in tastes. This leads to questions about the extent to which production tends to concentrate on a limited set of alternative bundles (see Rosen).

this information generally does not identify costs under a different overall level of attribute production.

This point may be most easily seen in Lucas (1975). Assume that there are more commodities than attributes and that each commodity contains a fixed level of attributes per unit. Then, in general, the solution for the attribute cost function, defined by Equation (13) above, will yield a set of convex (from above) piecewise linear iso-cost curves for attributes. However, if market demand shifts so that the overall consumption of commodities is shifted, then there will generally be a new set of commodity prices, p_t, and a new attribute cost function. In short, the attribute cost function has no more content than current market clearing prices for commodities.

Given unified markets, however, attribute costs may be expected to be the same from place to place, as long as commodity costs are also the same, markets are in equilibrium, and attributes are appropriately defined. Appropriate definition means in this case that the attribute content of commodities does not change. If, for example, an attribute is defined in terms of adequate air-conditioning or internal climate control, this definition is likely to involve quite different commodities in Pittsburgh and Phoenix, or even between the central city and rural-suburban areas of a given metropolitan area.

Apart from this, there is reason to believe that housing markets may be quite segmented both within and among different areas. First, the housing stock is extremely durable; given this durability and the high cost of transforming many attributes of the housing bundles, substantial heterogeneity of supply is likely. Thus market clearing prices for housing and housing attributes may vary from long-run equilibrium (and hence from place to place) over long periods.

Second, neighborhood amenities and public services are produced collectively, not competitively, and, like housing stock, they are not easily alterable in the short run. To the extent that households must choose from among a finite set of discrete attribute bundles, correlations between available housing attribute bundles and public services will mean that the "market" prices of unit attributes and public services will be determined jointly. In this case, unit attribute prices may vary from place to place and over time with variations in the association between unit types and public services.

Third, the housing market may be substantially segmented in ways that prevent the achievement of any unified long-run equilibrium position. For example, limited household mobility due to place attachment, strong ethnic ties, or racial discrimination may effectively divide a market into distinct submarkets. In the long run, each submarket might move toward a common equilibrium set of production costs; in the short run, however, market clearing prices may vary across submarkets.

The notion of locational variation in attribute prices is similar to the concept of quasi-rents, as discussed by Kain and Quigley (1975). Quasi-rents are formally described by:

(14) $$R_j = (P_1 + r_{1j})k_{1j} + (P_2 + r_{2j})k_{2j}, \ldots + (P_n + r_{nj})k_{nj} \quad j = 1 \ldots m,$$

where

R_j = housing price at the j^{th} location

P_i = supply price of i^{th} attribute if it is currently being produced, or market price at the least-cost location if it is not currently being produced

r_{ij} = quasi-rent for the i^{th} attribute at the j^{th} location

k_{ij} = quantity of i^{th} attribute at the j^{th} location.

The implicit price of each attribute consists of two components--the market value of each component, and the spatial quasi-rent for each attribute at the i^{th} location. The quasi-rents may be negative (for example, due to a decline in demand for the attribute after which the stock may not be profitably transformed), or positive (for example, due to a supply constraint, a spatial difference in production costs, or the presence of housing attributes that are not supplied by competitive firms).

The issue becomes more complex when joint costs are introduced. In principle, jointness can be handled by interactions between attribute prices and quasi-rents. Ultimately, however, this procedure leads to the definition of discrete housing types by specific values for only a few attributes. If many attribute types are considered, this analysis becomes unwieldy. For example, using 120 different rental bundle types and 12 neighborhood types, Apgar and Kain (1973) estimated 1,440 (12 x 120) different bundle prices.[1]

In summary, the hedonic cost function essentially provides no more information about tastes or production costs than does the set of commodity clearing prices. Hedonic cost functions for housing may be expected to be stable over time and from place to place to the extent that attributes are defined so that their commodity content is constant, and to the extent that the commodity prices in possibly segmented markets dominated by a substantial fixed stock do not vary from market to market.

The second price index issue involves the construction of aggregate indices. This issue does not arise in the application of hedonic indices within the Demand Experiment, since the housing index is computed for each individual; nevertheless, it is worth commenting on. The problem of aggregation is not peculiar to hedonic indices. Say, for example, that every individual receives a 10 percent increase in real income, and at the same time the price of some commodity (say meat) rises 10 percent (all other prices constant). It should be evident that the percentage increase in a household's real income (or alternatively, the true price index) will vary depending (approximately) on the share of income that the household spends on meat. Furthermore, if the budget share of meat is related to income or to other demographic variables, then different indices will apply to these groups. The use of hedonic indices does not change this situation (Muellbauer, 1974).

In addition, as discussed earlier, the hedonic cost function may be, and often is, estimated as a nonlinear form. This poses no special problems in computing individual indices: the Lespeyres index is simply computed

---

[1] If the hedonic form is misspecified, however, differences in estimated coefficients may reflect misspecification rather than the absence of long-run equilibrium. This point is emphasized by Muellbauer (1974) and Murray (forthcoming). Thus findings of price differences between black and white neighborhoods may reflect a misspecified market cost function rather than actual cost differences. This problem suggests that price differences must be tested for comparable bundles in order to establish discriminatory prices. As discussed in Section 1, however, the ultimate issue is access to different submarkets. Thus another approach to the effects of discrimination is to investigate the extent to which the range of accessible housing attribute bundles is limited for racial or ethnic minorities (see Merrill, 1976).

using the base period cost function. Likewise, an aggregate index may be computed by using mean attribute consumption (subject to the reservation that this applies only to households that consume the mean bundle).

Adelman and Griliches (1961) propose another index, based on marginal costs. If the attribute cost function is homogeneous of degree one in attributes, then the sum of the attributes, weighted by their marginal costs, will of course be the cost of the attributes (Muellbauer).[1] This is important to GNP deflaters when the production surface is not known, and when information is available only on product levels and marginal costs (prices) at a specific point. As Rosen points out, however, there is no reason to believe that the attribute cost function is homogeneous of degree one. It is true that the individual can (supposedly) buy more commodity units at fixed prices, so that some definition of attributes must exist for which the cost function is homogeneous. On the other hand, this definition is not necessarily the one used. As Rosen (1974) points out, two six-foot cars do not make a twelve-foot car (nor do two 150-square-foot rooms make a 300-square-foot room).

In conclusion, the present analysis is concerned with using hedonic regressions to derive estimates of housing quality. Evaluation of this particular use of hedonic equations rests on the ability to interpret the estimated coefficient in some consistent manner. The discussion above indicates that estimated coefficients are not likely to identify consumer preferences or underlying supply costs. A consumer preference or supply cost interpretation is not required, however. As long as the coefficients are reasonable approximations of the current market cost structure, the hedonic approach will provide consistent estimates of quality. If attribute prices vary widely in different submarkets, however, consistent interpretation of coefficients from a market-wide regression is open to question.

---

[1] This is a basic property of homogeneous functions of the first degree (Euler's Theorem).

REFERENCES

Adelman, Irma and Zvi Griliches, "On an Index of Quality Change," Journal of the American Statistical Association, vol. 56, September 1961, pp. 535-48.

Apgar, William and John Kain, Neighborhood Attributes and the Residential Price Geography of Urban Areas, Cambridge, Mass., National Bureau of Economic Research, 1973.

Court, Andrew, "Hedonic Price Indexes with Automotive Examples," Dynamics of Automobile Demand, New York, General Motors Corporation, 1939, pp. 99-117.

Griliches, Zvi, "Hedonic Price Indices Revisited," Price Indexes and Quality Change, Cambridge, Mass., Harvard University Press, 1971, pp. 3-15.

Kain, John and John Quigley, Housing Markets and Racial Discrimination, New York, National Bureau of Economic Research, 1975.

Lancaster, Kelvin, "A New Approach to Consumer Theory," Journal of Political Economy, vol. 74, April 1966, pp. 699-708.

Lucas, Robert E. B., "Hedonic Price Functions," Economic Inquiry, vol. 13, no. 2, June 1975.

Merrill, Sally R., The Effects of Segregation and Discrimination on the Price and Quality of Minority Housing, Ph.D. Dissertation, Boston University, 1976.

Muellbauer, John, "Household Production Theory, Quality, and the Hedonic Technique," American Economic Review, vol. 64, no. 6, December 1974, pp. 977-94.

Murray, Michael, "Hedonic Prices and Composite Commodities," Journal of Urban Economics (forthcoming).

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

Pollak, Robert A., "Subindexes in the Cost of Living Index," International Economic Review, vol. 16, no. 1, February 1976, pp. 135-150.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation Pure Competition," Journal of Political Economy, 1974, pp. 34-5.

Samuelson, Paul A., Foundations of Economic Analysis, Cambridge, Mass., Harvard University Press, 1947.

Triplett, Jack, "The Theory of Hedonic Quality Measurement and Its Use in Price Indexes," Bureau of Labor Statistics Staff Paper No. 6, Washington, D.C., Bureau of Labor Statistics, U.S. Department of Labor, 1971.

APPENDIX II

ESTIMATED HEDONIC EQUATIONS

Appendix II lists the hedonic equations for the common site variables; for submarkets, including central city and suburbs, white and minority submarkets, and for white and minority household subsamples; and for control households at enrollment and two years. These equations are discussed in Chapter 4.

LIST OF TABLES

LIST OF TABLES (continued)

Table II-1

PITTSBURGH LINEAR EQUATION COMMON SITE VARIABLES

$R^2 = 0.580$ $\bar{R}^2 = 0.576$ $F = 122.908$ $N = 1,618$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Length of residence (natural log)	-4.732	9.555
	Related to landlord (0,1)	-11.491	5.424
	Number of landlord contacts for maintenance	1.172	2.959
	Number of persons per room	6.814	4.063
Dwelling Unit Features	Area per room (natural log)	19.497	6.231
	Total number of rooms (includes kitchen & bath) (natural log)	53 452	27.878
	Building age (years)	-0.251	5.110
	Dishwasher and/or disposal provided (0,1)	11.042	4.472
	Adequate light and ventilation (0,1)	3.615	3.094
	Average surface and structural quality (4 point scale)	11.741	6.371
	Parking facilities provided (0,1)	8.614	5.426
	Large multifamily structure (0,1)	10.298	6.188
	Recent interior painting and papering (0,1)	7.433	4.107
	Many high quality features (0,1)	14.500	4.975
	Presence of private yard (0,1)	1.903	1 582
Neighborhood Features	Median income of census tract (dollars)	0.002	6.445
	Quality of block face landscaping (4 point scale)	4.502	7.025
	Quality of adult recreation facilities	17.176	8.985
CONSTANT		-133.064	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-2

PITTSBURGH SEMILOG EQUATION COMMON SITE VARIABLES

R^2 = 0 621 $\bar{R}^2$ = 0 616 F = 113 539 N = 1,617

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Charac-teristics	Length of residence (natural log)	- 044	10 020
	Related to landlord (0,1)	- 105	5.575
	Number of landlord contacts for maintenance	.012	3.396
	Number of persons per room	.083	5.602
Dwelling Unit Features	Area per room (natural log)	.189	6.808
	Total number of rooms (includes kitchen & bath)(natural log)	569	32.354
	Building age (years)	-.002	3 935
	Dishwasher and/or disposal provided (0,1)	052	2 382
	Adequate light and ventilation (0,1)	.025	2.422
	Average surface and structural quality (4 point scale)	087	5 170
	Parking facilities provided (0,1)	.058	4 144
	Large multifamily structure (0,1)	.062	3 954
	Recent interior painting or papering (0,1)	.050	3.129
	Many high quality features (0,1)	.071	2.731
	Presence of private yard (0,1)	N/A	N/A
	Temperature control· central heat or air-conditioning (0,1)	.018	1.493
	Inferior or no heat (0,1)	- 090	6.087
	Presence of adequate ceiling height (0,1)	.044	2 598
	Adequate plumbing present and working (0,1)	.020	1.444
	Stove and refrigerator provided (0,1)	.103	5.911
Neighborhood Features	Median income of census tract (dollars)	0	4 966
	Quality of block face landscaping (4 point scale)	.030	5 163
	Quality of adult recreation facilities	140	7 611
	Distance from Central Business District (miles)	-.002	1.104
CONSTANT		2.304	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

Table II-3

PHOENIX LINEAR EQUATION COMMON SITE VARIABLES

		R^2 = 0 771 $\bar{R}^2$ = 0.768 F = 296.502 N = 1,607	
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Length of residence (natural log)	-7 556	12.783
	Related to landlord (0,1)	-15.886	6.506
	Number of landlord contacts for maintenance	1.540	3.729
	Number of persons per room	7 018	5.232
Dwelling Unit Features	Area per room (natural log)	38.143	12.804
	Total number of rooms (includes kitchen & bath) (natural log)	77.940	32.667
	Building age (years)	-0.304	5 283
	Dishwasher and/or disposal provided (0,1)	11.000	5.825
	Adequate light and ventilation (0,1)	7.720	6.020
	Average surface and structural quality (4 point scale)	16.298	10.645
	Parking facilities provided (0,1)	5.986	4.518
	Large multifamily structure (0,1)	6.963	3.764
	Recent interior painting and papering (0,1)	3.382	2.332
	Many high quality features (0,1)	6.860	3.373
	Presence of private yard (0,1)	1.915	1.504
Neighborhood Features	Median income of census tract (dollars)	0.003	7.418
	Quality of block face landscaping (4 point scale)	2 943	4 098
	Quality of adult recreation facilities	11.937	4.941
CONSTANT		-241.526	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of population

Table II-4

PHOENIX SEMILOG EQUATION COMMON SITE VARIABLES

$R^2 = 0.786$ $\bar{R}^2 = 0.782$ $F = 252.036$ $N = 1{,}607$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Length of residence (natural log)	-.061	12.754
	Related to landlord (0,1)	-.132	6.769
	Number of landlord contacts for maintenance	.017	5.158
	Number of persons per room	.063	5.858
Dwelling Unit Features	Area per room (natural log)	.321	13.045
	Total number of rooms (includes kitchen & bath)(natural log)	.676	34.776
	Building age (years)	-.002	5.659
	Dishwasher and/or disposal provided (0,1)	.033	2.183
	Adequate light and ventilation (0,1)	.029	2.827
	Average surface and structural quality (4 point scale)	.127	9.205
	Parking facilities provided (0,1)	.037	3.505
	Large multifamily structure (0,1)	.030	2.030
	Recent interior painting or papering (0,1)	.020	1.718
	Many high quality features (0,1)	.021	1.246
	Presence of private yard (0,1)	N/A	N/A
	Temperature control. central heat or air-conditioning (0,1)	.060	4.426
	Inferior or no heat (0,1)	-.038	2.850
	Presence of adequate ceiling height (0,1)	.023	1.416
	Adequate plumbing present and working (0,1)	.025	1.758
	Stove and refrigerator provided (0,1)	.018	1.645
Neighborhood Features	Median income of census tract (dollars)	0	6.457
	Quality of block face landscaping (4 point scale)	.019	3.321
	Quality of adult recreation facilities	.083	4.244
	Distance from Central Business District (miles)	-.004	3.152
CONSTANT		1.718	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-5

POOLED SAMPLE SEMILOG INTERACTION EQUATION· COMMON SITE VARIABLES

R^2 = 0.734 $\bar{R}^2$ = 0.730 F = 187 673 N = 3,245

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Charac- teristics	Length of residence (natural log)	- 044	10 421
	Related to landlord (0,1)	- 103	5 650
	Number of landlord contacts for maintenance	012	3.536
	Number of persons per room	084	5.792
Dwelling Unit Features	Area per room (natural log)	190	7 056
	Total number of rooms (includes kitchen & bath) (natural log)	570	33.201
	Building age (years)	-.002	4.169
	Dishwasher and/or disposal provided (0,1)	.052	2.416
	Adequate light and ventilation (0,1)	026	2.570
	Average surface and structural quality (4 point scale)	083	5.054
	Parking facilities provided (0,1)	058	4.265
	Large multifamily structure (0,1)	063	4.160
	Recent interior painting or papering (0,1)	054	3.484
	Many high quality features (0,1)	071	2 788
	Temperature control central heat or air-conditioning (0,1)	019	1 673
	Inferior or no heat (0,1)	- 090	6.296
	Presence of adequate ceiling height (0,1)	.051	3.126
	Adequate plumbing present and working (0,1)	022	1 570
	Stove and refrigerator provided (0,1)	100	5 904
Neighborhood Features	Median income of census tract (dollars)	0	4 859
	Quality of block face landscaping (4 point scale)	029	5.114
	Quality of adult recreation facilities	141	7.829
	Distance from the Central Business District (miles)	-.001	0.840
Interaction, Tenure Characteristics	Interaction, length of residence (natural log) x Phoenix	-.017	2.599
	Interaction, related to landlord x Phoenix (0,1)	- 030	1.119
	Interaction, number of contacts for maintenance x Phoenix	005	1.019
	Interaction, number of persons per room x Phoenix	- 021	1 172
Interaction, Dwelling Unit Features	Interaction, area per room (natural log) x Phoenix	131	3.525
	Interaction, total number of rooms (includes kitchen and bath)(natural log) x Phoenix	107	4.059
	Interaction, building age (years) x Phoenix	- 001	1.384
	Interaction, dishwasher and/or disposal provided x Phoenix (0,1)	- 018	0.694
	Interaction, adequate light and ventilation x Phoenix (0,1)	003	0 212
	Interaction, average surface and structural quality (4 point scale) x Phoenix	044	2.020
	Interaction, parking facilities provided x Phoenix	- 021	1.210
	Interaction, large multifamily structure x Phoenix (0,1)	- 034	1.566
	Interaction, recent interior painting or papering x Phoenix (0,1)	- 034	1.744
	Interaction, many high quality features x Phoenix (0,1)	- 050	1.615
	Interaction, temperature control. central heat or air-conditioning x Phoenix (0,1)	041	2.264
	Interaction, inferior or no heat x Phoenix (0,1)	052	2.611
	Interaction, presence of adequate ceiling height x Phoenix (0,1)	-.027	1 171
	Interaction, adequate plumbing present and working x Phoenix	003	152
	Interaction, stove and refrigerator provided x Phoenix (0,1)	- 081	3 962
Interaction, Neighborhood Features	Interaction, median income of census tract (dollars x Phoenix	0	0.392
	Interaction, quality of block face landscaping (4 point scale) x Phoenix	- 010	1.204
	Interaction, quality of adult recreation facilities x Phoenix	- 058	2.138
	Interaction, distance from the Central Business District (miles) x Phoenix	- 002	1.296
CONSTANT		2.307	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-6

CENTRAL CITY LINEAR EQUATION PITTSBURGH

R^2 = 0 636 $\bar{R}^2$ = 0 622 F = 43.077 N = 820

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-7.893	2.909
	Length of residence (exponential function)	-14.893	8.136
	Landlord lives in the building (0,1)	-4.060	1.787
	Number of persons per room	7.772	3.658
	Number of landlord contacts for maintenance	0.822	1.751
Dwelling Unit Features	Area per room (natural log)	15.733	4.026
	Total number of rooms (includes kitchen & bath) (natural log)	58.638	22.384
	Building age (years)	-0 156	2.507
	Stove and refrigerator provided (0,1)	13.298	5.143
	Inferior or no heat (0,1)	-5.975	3.570
	Garage provided (0,1)	16.073	4.909
	Offstreet parking provided (0,1)	-1 467	0.510
	Overall evaluator rating (4 point scale)	4 613	3.425
	Dishwasher and/or disposal provided (0,1)	9.650	2 916
	Recent interior painting or papering (0,1)	5.321	2.361
	Many high quality features (0,1)	10.742	2.727
	Poor wall and ceiling surface (factor score)	-1.045	1.412
	Poor window condition (factor score)	-1.386	1.870
	Poor bathroom wall and ceiling surface (factor score)	-2 224	3.467
	High quality kitchen (0,1)	4.876	1.697
	Presence of adequate exits (0,1)	5.003	1.980
	Air-conditioning present (0,1)	1.525	0.662
	Presence of adequate ceiling height (0,1)	2.263	0 984
	Adequate kitchen facilities present (0,1)	5.420	0.696
	Large multifamily structure (0,1)	6.278	2 697
Neighborhood Features	Good recreational facilities and access (factor score)	2.609	2.433
	Traffic and litter problems (factor score)	-0.282	0.279
	Problems with crime and public services (factor score)	-1.313	1.604
	Census tracts with higher priced units and higher socioeconomic status (factor score)	3 214	3.253
	Nonminority census tracts with higher socio-economic status (factor score)	3.856	4.402
	Blue collar workers and nonminority residents in census tracts (factor score)	-2.973	4.527
	High quality block face (0,1)	4.478	2.776
CONSTANT		-81.289	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-7

CENTRAL CITY SEMILOG EQUATION: PITTSBURGH

R^2 = 0 646 $\bar{R}^2$ = 0 631 F = 41 806 N = 813

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	- 071	2.836
	Length of residence (exponential function)	-.144	8.539
	Landlord lives in the building (0,1)	-.042	2.000
	Number of persons per room	.085	4.330
	Number of landlord contacts for maintenance	.010	2.154
Dwelling Unit Features	Area per room (natural log)	.142	3.900
	Total number of rooms (includes kitchen & bath) (natural log)	562	22 520
	Building age (years)	-.001	2.402
	Stove and refrigerator provided (0,1)	.114	4.708
	Inferior or no heat (0,1)	- 062	3.978
	Garage provided (0,1)	.096	3.150
	Offstreet parking provided (0,1)	0	0
	Overall evaluator rating (4 point scale)	.049	3.928
	Dishwasher and/or disposal provided (0,1)	.063	2.072
	Recent interior painting or papering (0,1)	.043	2.052
	Many high quality features (0,1)	.055	1.518
	Poor wall and ceiling surface (factor score)	-.012	1.705
	Poor window condition (factor score)	-.009	1.341
	Poor bathroom wall and ceiling surface (factor score)	-.020	3.416
	High quality kitchen (0,1)	028	1.052
	Presence of adequate exits (0,1)	.052	2.224
	Air-conditioning present (0,1)	.021	1.008
	Presence of adequate ceiling height (0,1)	.026	1.208
	Adequate kitchen facilities present (0,1)	.142	1.964
	Large multifamily structure (0,1)	.064	2.927
	Working condition of plumbing (5 point scale)	.006	0.823
	Presence of private yard (0,1)	.021	1.530
Neighborhood Features	Good recreational facilities and access (factor score)	.022	2.228
	Traffic and litter problems (factor score)	-.003	0.338
	Problems with crime and public services (factor score)	-.016	2.105
	Census tracts with higher priced units and higher socioeconomic status (factor score)	.025	2.780
	Nonminority census tracts with higher socio-economic status (factor score)	.030	3.682
	Blue collar workers and nonminority residents in census tract (factor score)	-.028	4.577
	High quality block face (0,1)	.035	2.346
CONSTANT		2.734	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-8

SUBURBS LINEAR EQUATION· PITTSBURGH

R^2 = 0 678 $\bar{R}^2$ = 0.664 F = 49.171 N = 779

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-15.574	5.483
	Length of residence (exponential function)	-15.938	7.800
	Landlord lives in the building (0,1)	-6.764	2.578
	Number of persons per room	8.054	3.487
	Number of landlord contacts for maintenance	1.310	2.195
Dwelling Unit Features	Area per room (natural log)	22.393	5.053
	Total number of rooms (includes kitchen & bath) (natural log)	61.716	17.210
	Building age (years)	-0.271	4.024
	Stove and refrigerator provided (0,1)	15 572	5.295
	Inferior or no heat (0,1)	-9.481	4 184
	Garage provided (0,1)	13.097	4.776
	Offstreet parking provided (0,1)	4.688	1.943
	Overall evaluator rating (4 point scale)	5.605	3.712
	Dishwasher and/or disposal provided (0,1)	7.983	2.460
	Recent interior painting or papering (0,1)	7 595	3.072
	Many high quality features (0,1)	8 672	2.299
	Poor wall and ceiling surface (factor score)	-2.240	2.851
	Poor window condition (factor score)	-3.099	3.800
	Poor bathroom wall and ceiling surface (factor score)	-0.672	0.883
	High quality kitchen (0,1)	5.090	1.912
	Presence of adequate exits (0,1)	4.287	1.454
	Air-conditioning present (0,1)	4.425	1.845
	Presence of adequate ceiling height (0,1)	4.263	1.578
	Adequate kitchen facilities present (0,1)	7.781	0.929
	Large multifamily structure (0,1)	0.379	0 155
Neighborhood Features	Good recreational facilities and access (factor score)	2.290	2.905
	Traffic and litter problems (factor score)	-1.850	2.236
	Problems with crime and public services (factor score)	-1.603	1.726
	Census tracts with higher priced units and socioeconomic status (factor score)	3.501	3 910
	Nonminority census tracts with higher socio-economic status (factor score)	3.800	3.372
	Blue collar workers and nonminority residents in census tract (factor score)	-3.642	3.488
	High quality block face (0,1)	5.461	3.318
CONSTANT		-117.032	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-9

SUBURBS SEMILOG EQUATION PITTSBURGH

$R^2 = 0.686$ $\bar{R}^2 = 0.671$ $F = 41.119$ $N = 770$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.133	5.255
	Length of residence (exponential function)	-.145	8.079
	Landlord lives in the building (0,1)	-.094	4.069
	Number of persons per room	.087	4.310
	Number of landlord contacts for maintenance	.014	2.591
Dwelling Unit Features	Area per room (natural log)	.186	4.744
	Total number of rooms (includes kitchen & bath) (natural log)	.565	17.354
	Building age (years)	-.002	3.019
	Stove and refrigerator provided (0,1)	.098	3.819
	Inferior or no heat (0,1)	-.108	5.447
	Garage provided (0,1)	.085	3.493
	Offstreet parking provided (0,1)	.031	1.448
	Overall evaluator rating (4 point scale)	.055	4.110
	Dishwasher and/or disposal provided (0,1)	.046	1.626
	Recent interior painting or papering (0,1)	.063	2.891
	Many high quality features (0,1)	.031	0.952
	Poor wall and ceiling surface (factor score)	-.025	3.638
	Poor window condition (factor score)	-.027	3.744
	Poor bathroom wall and ceiling surface (factor score)	-.002	0.375
	High quality kitchen (0,1)	.032	1.382
	Presence of adequate exits (0,1)	.044	1.711
	Air-conditioning present (0,1)	.032	1.516
	Presence of adequate ceiling height (0,1)	.045	1.907
	Adequate kitchen facilities present (0,1)	.091	1.211
	Large multifamily structure (0,1)	.018	0.810
	Working condition of plumbing (5 point scale)	.009	1.136
	Presence of private yard (0,1)	.010	0.703
Neighborhood Features	Good recreational facilities and access (factor score)	.021	3.052
	Traffic and litter problems (factor score)	-.017	2.273
	Problems with crime and public services (factor score)	-.017	2.097
	Census tracts with higher priced units and higher socioeconomic status (factor score)	.033	4.183
	Nonminority census tracts with higher socio-economic status (factor score)	.042	4.226
	Blue collar workers and nonminority residents in census tract (factor score)	-.026	2.845
	High quality block face (0,1)	.042	2.922
CONSTANT		2.572	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

Table II-10

CENTRAL CITY LINEAR EQUATION. PHOENIX

	R^2 = 0 785 $\bar{R}^2$ = 0 781 F = 181.821 N = 1,217		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-14.517	5.266
	Length of residence (exponential function)	-22.835	10.954
	Number of persons per room	8.602	5.717
	Number of landlord contacts for maintenance	1.240	2.834
Dwelling Unit Features	Area per room (natural log)	36.091	10.626
	Total number of rooms (includes kitchen & bath) (natural log)	79.238	29.339
	Building age (years)	-0.246	3.817
	Stove or refrigerator provided (0,1)	3.285	1.829
	Central heat present (0,1)	7.321	3.655
	Garage or carport provided (0,1)	3.969	2.761
	Dishwasher and/or disposal provided (0,1)	6.905	3.148
	Recent interior painting or papering (0,1)	3 180	2.058
	Average surface and structural quality (4 point scale)	13.582	7 834
	Adequate light and ventilation (0,1)	6.006	4 280
	Central air-conditioning present (0,1)	7.559	3.190
	Large multifamily structure (0,1)	4.171	2.012
Neighborhood Features	Overall neighborhood quality (factor score)	2.723	2.944
	Recreational facilities (factor score)	0.333	0.356
	Access to shopping and parking (factor score)	-1.330	1.415
	Census tracts with higher priced dwelling units and higher socioeconomic status (factor score)	4.799	3.802
	Owner-occupied, single-family units in census tract (factor score)	1.215	1.435
	Poor quality housing in census tract (factor score)	-3.085	3.825
	Distance from the Central Business District (miles)	0.703	1.739
	Quality of block face landscaping (4 point scale)	3.236	4.007
CONSTANT		-209.777	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

Table II-11

CENTRAL CITY SEMILOG EQUATION PHOENIX

R^2 = 0.802 $\bar{R}^2$ = 0 798 F = 178 672 N = 1,217

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.139	6.320
	Length of residence (exponential function)	-.196	11.846
	Number of persons per room	.079	6.597
	Number of landlord contacts for maintenance	.014	4.095
Dwelling Unit Features	Area per room (natural log)	.306	11,172
	Total number of rooms (includes kitchen & bath) (natural log)	.683	30.676
	Building age(years)	-.002	4.641
	Stove or refrigerator provided (0,1)	.023	1.634
	Central heat present (0,1)	.029	1.788
	Garage or carport provided (0,1)	.027	2.355
	Dishwasher and/or disposal provided (0,1)	.027	1.577
	Recent interior painting or papering (0,1)	.021	1.738
	Average surface and structural quality (4 point scale)	.123	8.118
	Adequate light and ventilation (0,1)	034	3.030
	Central air-conditioning present (0,1)	.064	3.389
	Large multifamily structure (0,1)	.021	I 272
	Plumbing present (0,1)	.062	2.873
	Inferior or no heat (0,1)	-.016	1.094
	Presence of adequate ceiling height (0,1)	.007	0 373
Neighborhood Features	Overall neighborhood quality (factor score)	.025	3.384
	Recreational facilities (factor score)	.008	1.134
	Access to shopping and parking (factor score)	.006	0.738
	Census tracts with higher priced units and higher socioeconomic status (factor score)	.030	2.969
	Owner-occupied single family dwelling units in census tract (factor score)	007	0.980
	Poor quality housing in census tract (factor score)	-.033	5.167
	Distance from the Central Business District (miles)	-.001	0.354
	Quality of block face landscaping (4 point scale)	025	3.849
CONSTANT		1.885	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-12

SUBURBS LINEAR EQUATION PHOENIX

R^2 = 0 810 $\bar{R}^2$ = 0 798 F = 62 700 N = 377

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-15.303	3.476
	Length of residence (exponential function)	-22.269	5 555
	Number of persons per room	4 438	1.718
	Number of landlord contacts for maintenance	1.116	1 237
Dwelling Unit Features	Area per room (natural log)	36.633	5.982
	Total number of rooms (includes kitchen & bath) (natural log)	78 353	14.456
	Building age (years)	-0.213	1.659
	Stove or refrigerator provided (0,1)	8.160	2.332
	Central heat present (0,1)	9 338	2.361
	Garage or carport provided (0,1)	7.106	2.693
	Dishwasher and/or disposal provided (0,1)	12.698	3.656
	Recent interior painting or papering (0,1)	-1.528	0.480
	Average surface and structural quality (4 point scale)	13.165	3.909
	Adequate light and ventilation (0,1)	7.134	2.753
	Central air-conditioning present (0,1)	6.407	1.576
	Large multifamily structure (0,1)	7.636	2.126
Neighborhood Features	Overall neighborhood quality (factor score)	0.324	0.197
	Recreational facilities (factor score)	2.030	1.705
	Access to shopping and parking (factor score)	9.683	2.815
	Census tracts with higher priced dwelling units and higher socioeconomic status (factor score)	0.343	0.176
	Owner-occupied, single-family units in census tract (factor score)	2.690	1.694
	Poor quality housing in census tract (factor score)	-2 884	1.753
	Distance from the Central Business District (miles)	-0.342	0 724
	Quality of block face landscaping (4 point scale)	2.115	1.445
CONSTANT		-209.984	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-13

SUBURBS SEMILOG EQUATION PHOENIX

R^2 = 0 827 $\bar{R}^2$ = 0 813 F = 61 671 N = 377

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.107	3.173
	Length of residence (exponential function)	-.188	6.094
	Number of persons per room	.025	1.262
	Number of landlord contacts for maintenance	.015	2.135
Dwelling Unit Features	Area per room (natural log)	.314	6.513
	Total number of rooms (includes kitchen & bath) (natural log)	.623	14.026
	Building age(years)	-.001	1.318
	Stove or refrigerator provided (0,1)	.055	2.044
	Central heat present (0,1)	.066	2.171
	Garage or carport provided (0,1)	.046	2.301
	Dishwasher and/or disposal provided (0,1)	.060	2.250
	Recent interior painting or papering (0,1)	- 010	0.411
	Average surface and structural quality (4 point scale)	.111	4.065
	Adequate light and ventilation (0,1)	.027	1.347
	Central air-conditioning present (0,1)	024	0.774
	Large multifamily structure (0,1)	.046	1.686
	Plumbing present (0,1)	006	0 158
	Inferior or no heat (0,1)	-.102	3.346
	Presence of adequate ceiling height (0,1)	.059	2.002
Neighborhood Features	Overall neighborhood quality (factor score)	-.005	0.406
	Recreational facilities (factor score)	.011	1.199
	Access to shopping and parking (factor score)	072	2.729
	Census tracts with higher priced units and higher socioeconomic status (factor score)	-.004	0.251
	Owner-occupied single family dwelling units in census tract (factor score)	.016	1.308
	Poor quality housing in census tract (factor score)	-.024	1.801
	Distance from the Central Business District (miles)	.0001	0.032
	Quality of block face landscaping (4 point scale)	.017	1.480
CONSTANT		1.956	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-14

GHETTO SUBMARKET LINEAR EQUATION: PITTSBURGH

R^2 = 0.670 $\bar{R}^2$ = 0 621 F = 13 703 N = 241

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-6.360	1.121
	Length of residence (exponential function)	-9.696	3.122
	Landlord lives in the building (0,1)	-6.439	1.156
	Number of persons per room	1.992	0.619
	Number of landlord contacts for maintenance	1.559	2 250
Dwelling Unit Features	Area per room (natural log)	16.885	2.329
	Total number of rooms (includes kitchen & bath) (natural log)	52.196	11.129
	Building age (years)	-0 081	0 661
	Stove and refrigerator provided (0,1)	15.952	3.021
	Inferior or no heat (0,1)	-6.552	2.315
	Garage provided (0,1)	9.062	1.822
	Offstreet parking provided (0,1)	2.515	0.522
	Overall evaluator rating (4 point scale)	5.268	2.330
	Dishwasher and/or disposal provided (0,1)	14.232	2.254
	Recent interior painting or papering (0,1)	11 441	2.833
	Many high quality features (0,1)	18.906	2.199
	Poor wall and ceiling surface (factor score)	1.988	1.509
	Poor window condition (factor score)	0.213	0.167
	Poor bathroom wall and ceiling surface (factor score)	-1.324	1.194
	High quality kitchen (0,1)	-1.466	0.253
	Presence of adequate exits (0,1)	11.747	2.226
	Air-conditioning present (0,1)	4 222	0.839
	Presence of adequate ceiling height (0,1)	2.066	0.562
	Adequate kitchen facilities present (0,1)	9 078	1.129
	Large multifamily structure (0,1)	8.910	1.749
Neighborhood Features	Good recreational facilities and access (factor score)	4.050	2.170
	Traffic and litter problems (factor score)	-2.240	0.855
	Problems with crime and public services (factor score)	1.137	0.365
	Census tracts with higher socioeconomic status (factor score)	4.164	2.012
	Census tracts with newer, higher priced units (factor score)	1.072	0 532
	High quality block face (0,1)	8 368	2 690
CONSTANT		-91.320	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

Table II-15

GHETTO SUBMARKET SEMILOG EQUATION   PITTSBURGH

$R^2 = 0.696$ $\bar{R}^2 = 0.647$ $F = 14.207$ $N = 239$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.047	0.854
	Length of residence (exponential function)	-.093	3.047
	Landlord lives in the building (0,1)	-.062	1.124
	Number of persons per room	.045	1.439
	Number of landlord contacts for maintenance	.018	2.484
Dwelling Unit Features	Area per room (natural log)	.143	2.004
	Total number of rooms (includes kitchen & bath) (natural log)	.553	11.186
	Building age (years)	0	0 251
	Stove and refrigerator provided (0,1)	.088	1 706
	Inferior or no heat (0,1)	-.069	2.473
	Garage provided (0,1)	.053	1.104
	Offstreet parking provided (0,1)	.026	0.560
	Overall evaluator rating (4 point scale)	.064	2 826
	Dishwasher and/or disposal provided (0,1)	.101	1.626
	Recent interior painting or papering (0,1)	118	3 013
	Many high quality features (0,1)	.075	0.894
	Poor wall and ceiling surface (factor score)	.016	1.257
	Poor window condition (factor score)	-.003	0.251
	Poor bathroom wall and ceiling surface (factor score)	-.014	1.326
	High quality kitchen (0,1)	-.006	0.114
	Presence of adequate exits (0,1)	.129	2.504
	Air-conditioning present (0,1)	.049	1.005
	Presence of adequate ceiling height (0,1)	.023	0.655
	Adequate kitchen facilities present (0,1)	.154	1.928
	Large multifamily structure (0,1)	.092	1.838
	Working condition of plumbing (5 point scale)	-.008	0.621
	Presence of private yard (0,1)	.024	0.928
Neighborhood Features	Good recreational facilities and access (factor score)	.052	2.859
	Traffic and litter problems (factor score)	-.024	0.922
	Problems with crime and public services (factor score)	.026	0.862
	Census tracts with higher socioeconomic status (factor score)	.033	1.617
	Census tracts with newer, higher priced units (factor score)	.014	0.711
	High quality block face (0,1)	.076	2.486
CONSTANT		2.599	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-16

WHITE SUBMARKET LINEAR EQUATION PITTSBURGH

$R^2 = 0.666$ $\quad \bar{R}^2 = 0.657$ $\quad F = 74.637$ $\quad N = 1,192$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-13.369	5.762
	Length of residence (exponential function)	-16.133	9.842
	Landlord lives in the building (0,1)	-5.379	2.751
	Number of persons per room	8 578	4 472
	Number of landlord contacts for maintenance	0 909	1.950
Dwelling Unit Features	Area per room (natural log)	21.038	6.136
	Total number of rooms (includes kitchen & bath) (natural log)	62.814	24.722
	Building age (years)	-0.262	5.059
	Stove and refrigerator provided (0,1)	14.863	6.682
	Inferior or no heat (0,1)	-7.825	4.671
	Garage provided (0,1)	14.777	6 158
	Offstreet parking provided (0,1)	3.193	1.566
	Overall evaluator rating (4 point scale)	5.167	4 275
	Dishwasher and/or disposal provided (0,1)	8.185	3.029
	Recent interior painting or papering (0,1)	5.172	2.661
	Many high quality features (0,1)	10.126	3.333
	Poor wall and ceiling surface (factor score)	-1.997	3.130
	Poor window condition (factor score)	-2.808	4.306
	Poor bathroom wall and ceiling surface (factor score)	-1.783	3.010
	High quality kitchen (0,1)	4.637	2.154
	Presence of adequate exits (0,1)	4.982	2.301
	Air-conditioning present (0,1)	2.296	1.226
	Presence of adequate ceiling height (0,1)	3.269	1.504
	Adequate kitchen facilities present (0,1)	4.716	0.572
	Large multifamily structure (0,1)	3 037	1 616
Neighborhood Features	Good recreational facilities and access (factor score)	2.518	4.168
	Traffic and litter problems (factor score)	-0.986	1.404
	Problems with crime and public services (factor score)	-1.626	2.530
	Census tracts with higher socioeconomic status (factor score)	6.392	8.413
	Census tracts with newer, higher priced units (factor score)	2.316	3.103
	High quality block face (0,1)	3.828	2.888
CONSTANT		-108.752	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-17

WHITE SUBMARKET SEMILOG EQUATION PITTSBURGH

R^2 = 0.668 $\bar{R}^2$ = 0.659 F = 69 980 N = 1,180

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	- 116	5.625
	Length of residence (exponential function)	-.147	10.110
	Landlord lives in the building (0,1)	-.069	3 996
	Number of persons per room	.087	5.112
	Number of landlord contacts for maintenance	.009	2.245
Dwelling Unit Features	Area per room (natural log)	181	5.925
	Total number of rooms (includes kitchen & bath) (natural log)	.573	24.712
	Building age (years)	-.002	4.138
	Stove and refrigerator provided (0,1)	.117	5.870
	Inferior or no heat (0,1)	-.088	5.923
	Garage provided (0,1)	.096	4.423
	Offstreet parking provided (0,1)	022	1.235
	Overall evaluator rating (4 point scale)	.051	4.738
	Dishwasher and/or disposal provided (0,1)	.040	1.685
	Recent interior painting or papering (0,1)	.037	2.164
	Many high quality features (0,1)	.048	1.802
	Poor wall and ceiling surface (factor score)	-.023	4.029
	Poor window condition (factor score)	-.021	3.653
	Poor bathroom wall and ceiling surface (factor score)	-.015	2.879
	High quality kitchen (0,1)	.028	1.442
	Presence of adequate exits (0,1)	.048	2.489
	Air-conditioning present (0,1)	.015	0.902
	Presence of adequate ceiling height (0,1)	.035	1.827
	Adequate kitchen facilities present (0,1)	.109	1.470
	Large multifamily structure (0,1)	.036	2.108
	Working condition of plumbing (5 point scale)	.009	1.517
	Presence of private yard (0,1)	.014	1.171
Neighborhood Features	Good recreational facilities and access (factor score)	.023	4.208
	Traffic and litter problems (factor score)	- 008	1.313
	Problems with crime and public services (factor score)	-.018	3.167
	Census tracts with higher socioeconomic status (factor score)	.056	8.140
	Census tracts with newer, higher priced units (factor score)	020	2.984
	High quality block face (0,1)	.029	2.451
CONSTANT		2.582	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-18

SPANISH AMERICAN SUBMARKET LINEAR EQUATION. PHOENIX

	R^2 = 0.693 $\bar{R}^2$ = 0.654 F = 17 780 N = 214		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-9.369	2.362
	Length of residence (exponential function)	-14.631	3.905
	Number of persons per room	2.334	1.259
	Number of landlord contacts for maintenance	1.612	2.101
Dwelling Unit Features	Area per room (natural log)	26.811	4.582
	Total number of rooms (includes kitchen & bath) (natural log)	48.944	10.845
	Building age (years)	-0.258	2.240
	Stove or refrigerator provided (0,1)	-3 974	1.457
	Central heat present (0,1)	-3.973	0.560
	Garage or carport provided (0,1)	5.516	1.627
	Dishwasher and/or disposal provided (0,1)	8.330	1.055
	Recent interior painting or papering (0,1)	7.943	1.988
	Average surface and structural quality (4 point scale)	12.978	5 106
	Adequate light and ventilation (0,1)	4 176	1.114
	Central air-conditioning present (0,1)	13.030	1.395
	Large multifamily structure (0,1)	11.149	1.536
Neighborhood Features	Overall neighborhood quality (factor score)	1.547	0.550
	Recreational facilities (factor score)	5.692	2.534
	Access to shopping and parking (factor score)	-2.555	0.894
	Census tracts with higher priced dwelling units and higher socioeconomic status (factor score)	11.184	1.580
	Owner-occupied, single-family units in census tract (factor score)	7.053	2.103
	Poor quality housing in census tract (factor score)	-6 722	2.797
	Distance from the Central Business District (miles)	0.029	0.063
	Quality of block face landscaping (4 point scale)	5.365	3.225
CONSTANT		-105 630	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-19

SPANISH AMERICAN SUBMARKET SEMILOG EQUATION PHOENIX

$R^2 = 0.732$ $\bar{R}^2 = 0.694$ $F = 19\ 606$ $N = 214$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.081	1.967
	Length of residence (exponential function)	- 149	3.893
	Number of persons per room	024	1 254
	Number of landlord contacts for maintenance	.019	2.363
Dwelling Unit Features	Area per room (natural log)	.323	5.279
	Total number of rooms (includes kitchen & bath) (natural log)	566	11.357
	Building age (years)	-.002	1.954
	Stove or refrigerator provided (0,1)	-.036	1 271
	Central heat present (0,1)	- 054	0.714
	Garage or carport provided (0,1)	.055	1.589
	Dishwasher and/or disposal provided (0,1)	029	0.352
	Recent interior painting or papering (0,1)	.094	2.298
	Average surface and structural quality (4 point scale)	.098	3.360
	Adequate light and ventilation (0,1)	.028	0.727
	Central air-conditioning present (0,1)	.067	0.693
	Large multifamily structure (0,1)	.147	1.971
	Plumbing present (0,1)	051	1.682
	Inferior or no heat (0,1)	-.069	1.821
	Presence of adequate ceiling height (0,1)	.012	0.390
Neighborhood Features	Overall neighborhood quality (factor score)	.013	0 476
	Recreational facilities (factor score)	.038	1.899
	Access to shopping and parking (factor score)	-.007	0.251
	Census tracts with higher priced units and higher socioeconomic status (factor score)	.085	1.219
	Poor quality housing in census tract (factor score)	-.051	2.221
	Distance from the Central Business District (miles)	004	0.919
	Quality of block face landscaping (4 point scale)	.049	2.857
CONSTANT		2.145	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-20

WHITE SUBMARKET LINEAR EQUATION. PHOENIX

	R^2 = 0 773 $\bar{R}^2$ = 0.767 F = 125.791 N = 912		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-21.710	5 764
	Length of residence (exponential function)	-22.984	8 769
	Number of persons per room	10.894	4.929
	Number of landlord contacts for maintenance	0.187	0.346
Dwelling Unit Features	Area per room (natural log)	45.622	11.033
	Total number of rooms (includes kitchen & bath) (natural log)	94.729	26.752
	Building age (years)	-0.330	4.009
	Stove or refrigerator provided (0,1)	6.903	2.750
	Central heat present (0,1)	5.704	2.699
	Garage or carport provided (0,1)	4.239	2.661
	Dishwasher and/or disposal provided (0,1)	9 332	4.347
	Recent interior painting or papering (0,1)	1.604	0.914
	Average surface and structural quality (4 point scale)	14.425	6.014
	Adequate light and ventilation (0,1)	6.540	4.257
	Central air-conditioning present (0,1)	7.128	3.081
	Large multifamily structure (0,1)	4.215	2.026
Neighborhood Features	Overall neighborhood quality (factor score)	3.669	3.536
	Recreational facilities (factor score)	2.316	2.719
	Access to shopping and parking (factor score)	1.742	1.876
	Census tracts with higher priced dwelling units and higher socioeconomic status (factor score)	2.498	1.645
	Owner-occupied, single-family units in census tract (factor score)	0.429	0.524
	Poor quality housing in census tract (factor score)	-0.724	0.847
	Distance from the Central Business District (miles)	-0.557	2.408
	Quality of block face landscaping (4 point scale)	0.904	1 003
CONSTANT		-273.000	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-21

WHITE SUBMARKET SEMILOG EQUATION PHOENIX

	R^2 = 0 778 $\bar{R}^2$ = 0 771 F = 119.074 N =912		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.152	5.417
	Length of residence (exponential function)	-.184	9.397
	Number of persons per room	.080	4.852
	Number of landlord contacts for maintenance	.005	1.320
Dwelling Unit Features	Area per room (natural log)	361	11.450
	Total number of rooms (includes kitchen & bath) (natural log)	.706	26.493
	Building age (years)	- 003	5.737
	Stove or refrigerator provided (0,1)	.060	3.215
	Central heat present (0,1)	.027	1 693
	Garage or carport provided (0,1)	.030	2.483
	Dishwasher and/or disposal provided (0,1)	.034	2.140
	Recent interior painting or papering (0,1)	.011	0.875
	Average surface and structural quality (4 point scale)	.095	4.941
	Adequate light and ventilation (0,1)	.037	3.204
	Central air-conditioning present (0,1)	.052	3.030
	Large multifamily structure (0,1)	.017	1.100
	Plumbing present (0,1)	.063	1.650
	Inferior or no heat (0,1)	-.036	1.993
	Presence of adequate ceiling height (0,1)	.032	1.363
Neighborhood Features	Overall neighborhood quality (factor score)	032	4.105
	Recreational facilities (factor score)	.013	2.078
	Access to shopping and parking (factor score)	.019	2.718
	Census tracts with higher priced units and higher socioeconomic status (factor score)	022	1.943
	Poor quality housing in census tract (factor score)	- 005	0 857
	Distance from the Central Business District (miles)	-.004	2.655
	Quality of block face landscaping (4 point scale)	.006	0.886
CONSTANT		1.692	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation, Form, 1970 Census of Population.

Table II-22

BLACK HOUSEHOLDS LINEAR EQUATION: PITTSBURGH

R^2 = 0 648 $\bar{R}^2$ = 0.616 F = 19 752 N = 364

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-6.343	1.764
	Length of residence (exponential function)	-9.515	3.920
	Landlord lives in the building (0,1)	-3.262	0.848
	Number of persons per room	1.551	0.622
	Number of landlord contacts for maintenance	1.035	1.764
Dwelling Unit Features	Area per room (natural log)	14.177	2.506
	Total number of rooms (includes kitchen & bath) (natural log)	52.106	13.462
	Building age (years)	-0 170	1.806
	Stove and refrigerator provided (0,1)	13.417	3.014
	Inferior or no heat (0,1)	-5.410	2.520
	Garage provided (0,1)	10.152	2.638
	Offstreet parking provided (0,1)	4.202	0.956
	Overall evaluator rating (4 point scale)	4.647	2.489
	Dishwasher and/or disposal provided (0,1)	9.430	2.170
	Recent interior painting or papering (0,1)	6.855	2.095
	Many high quality features (0,1)	22.682	2.833
	Poor wall and ceiling surface (factor score)	0.295	0.300
	Poor window condition (factor score)	-0.187	0.190
	Poor bathroom wall and ceiling surface (factor score)	-1.255	1 422
	High quality kitchen (0,1)	6.768	1.430
	Presence of adequate exits (0,1)	9.317	2.289
	Air-conditioning present (0,1)	1.620	0.463
	Presence of adequate ceiling height (0,1)	5.807	2.029
	Adequate kitchen facilities present (0,1)	5.301	0.592
	Large multifamily structure (0,1)	3.327	0.678
Neighborhood Features	Good recreational facilities and access (factor score)	3.112	2.791
	Traffic and litter problems (factor score)	-0.388	0.281
	Problems with crime and public services (factor score)	2.606	1.841
	Census tracts with higher socioeconomic status (factor score)	2.891	1.978
	Census tracts with newer, higher priced units (factor score)	-0.178	0.138
	High quality block face (0,1)	8.735	3.617
CONSTANT		-72.212	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-23

BLACK HOUSEHOLDS SEMILOG EQUATION. PITTSBURGH

$R^2 = 0.672$ $\bar{R}^2 = 0.638$ $F = 20.328$ $N = 362$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.038	1.126
	Length of residence (exponential function)	-.102	4.478
	Landlord lives in the building (0,1)	-.039	1.078
	Number of persons per room	.032	1.378
	Number of landlord contacts for maintenance	.014	2.359
Dwelling Unit Features	Area per room (natural log)	.089	1.643
	Total number of rooms (includes kitchen & bath) (natural log)	.518	13.303
	Building age (years)	-.001	1.485
	Stove and refrigerator provided (0,1)	.075	1.794
	Inferior or no heat (0,1)	-.060	2.938
	Garage provided (0,1)	.051	1.414
	Offstreet parking provided (0,1)	.046	1.099
	Overall evaluator rating (4 point scale)	.050	2.819
	Dishwasher and/or disposal provided (0,1)	.081	1.981
	Recent interior painting or papering (0,1)	.083	2.648
	Many high quality features (0,1)	.112	1.494
	Poor wall and ceiling surface (factor score)	.006	0.596
	Poor window condition (factor score)	-.004	0.443
	Poor bathroom wall and ceiling surface (factor score)	-.012	1.450
	High quality kitchen (0,1)	.040	0.910
	Presence of adequate exits (0,1)	.087	2.262
	Air-conditioning present (0,1)	.020	0.616
	Presence of adequate ceiling height (0,1)	.055	2.018
	Adequate kitchen facilities present (0,1)	.132	1.519
	Large multifamily structure (0,1)	.064	1.373
	Working condition of plumbing (5 point scale)	-.007	0.692
	Presence of private yard (0,1)	.042	2.154
Neighborhood Features	Good recreational facilities and access (factor score)	.036	3.473
	Traffic and litter problems (factor score)	.002	0.148
	Problems with crime and public services (factor score)	.024	1.773
	Census tracts with higher socioeconomic status (factor score)	.027	1.977
	Census tracts with newer, higher priced units (factor score)	-.003	0.247
	High quality block face (0,1)	.075	3.324
CONSTANT		3.016	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-24

WHITE HOUSEHOLDS LINEAR EQUATION. PITTSBURGH

	R^2 = 0.662 $\bar{R}^2$ = 0.653 F = 75 091 N = 1,219		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-13.469	5.803
	Length of residence (exponential function)	-16 841	10.392
	Landlord lives in the building (0,1)	-5.840	3.045
	Number of persons per room	9.802	5.062
	Number of landlord contacts for maintenance	1.090	2.341
Dwelling Unit Features	Area per room (natural log)	20.802	6.180
	Total number of rooms (includes kitchen & bath) (natural log)	62.051	25.054
	Building age (years)	-0.235	4.570
	Stove and refrigerator provided (0,1)	13.889	6.378
	Inferior or no heat (0,1)	-8.089	4.854
	Garage provided (0,1)	15.557	6.453
	Offstreet parking provided (0,1)	2.023	0.982
	Overall evaluator rating (4 point scale)	4 714	3.981
	Dishwasher and/or disposal provided (0,1)	9.268	3.471
	Recent interior painting or papering (0,1)	5.969	3.083
	Many high quality features (0,1)	6.522	2.215
	Poor wall and ceiling surface (factor score)	-2.317	3.685
	Poor window condition (factor score)	-3.038	4.701
	Poor bathroom wall and ceiling surface (factor score)	-1.575	2.701
	High quality kitchen (0,1)	5.458	2.535
	Presence of adequate exits (0,1)	3.536	1.626
	Air-conditioning present (0,1)	2.245	1.197
	Presence of adequate ceiling height (0,1)	1.603	.743
	Adequate kitchen facilities present (0,1)	7.239	1.022
	Large multifamily structure (0,1)	3.567	1.955
Neighborhood Features	Good recreational facilities and access (factor score)	2.515	4.162
	Traffic and litter problems (factor score)	-1.070	1.523
	Problems with crime and public services (factor score)	-1.521	2.385
	Census tracts with higher socioeconomic status (factor score)	6.333	8.561
	Census tracts with newer, higher priced units (factor score)	2.910	3.750
	High quality block face (0,1)	3.934	3.034
CONSTANT		-106.462	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-25

WHITE HOUSEHOLDS SEMILOG EQUATION: PITTSBURGH

	R^2 = 0.668 $\bar{R}^2$ = 0 658 F = 71.245 N = 1,205		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	- 119	5 708
	Length of residence (exponential function)	-.155	10.682
	Landlord lives in the building (0,1)	- 072	4.187
	Number of persons per room	.099	5.712
	Number of landlord contacts for maintenance	.012	2.779
Dwelling Unit Features	Area per room (natural log)	.188	6.216
	Total number of rooms (includes kitchen & bath) (natural log)	.570	25.059
	Building age (years)	-.002	3.693
	Stove and refrigerator provided (0,1)	.108	5 518
	Inferior or no heat (0,1)	-.091	6.099
	Garage provided (0,1)	.100	4.588
	Offstreet parking provided (0,1)	012	0.666
	Overall evaluator rating (4 point scale)	.048	4 472
	Dishwasher and/or disposal provided (0,1)	052	2.193
	Recent interior painting or papering (0,1)	.042	2.453
	Many high quality features (0,1)	.024	0.938
	Poor wall and ceiling surface (factor score)	-.027	4.830
	Poor window condition (factor score)	-.024	4.145
	Poor bathroom wall and ceiling surface (factor score)	-.012	2.316
	High quality kitchen (0,1)	.032	1.662
	Presence of adequate exits (0,1)	.036	1.829
	Air-conditioning present (0,1)	.015	0.877
	Presence of adequate ceiling height (0,1)	020	1.041
	Adequate kitchen facilities present (0,1)	.116	1.831
	Large multifamily structure (0,1)	.040	2.438
	Working condition of plumbing (5 point scale)	.013	2.110
	Presence of private yard (0,1)	.014	1.176
Neighborhood Features	Good recreational facilities and access (factor score)	.023	4.211
	Traffic and litter problems (factor score)	-.009	1.399
	Problems with crime and public services (factor score)	-.016	2.873
	Census tracts with higher socioeconomic status (factor score)	.054	8.104
	Census tracts with newer, higher priced units (factor score)	.025	3.652
	High quality block face (0,1)	.032	2.738
CONSTANT		2.554	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-26

SPANISH HOUSEHOLDS LINEAR EQUATION: PHOENIX

$R^2 = 0.776$ $\bar{R}^2 = 0.761$ $F = 51.056$ $N = 378$

VARIABLE DESCRIPTION		COEFFICIENTS	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-6.539	1.541
	Length of residence (exponential function)	-20.780	5.751
	Number of persons per room	5.763	2.852
	Number of landlord contacts for maintenance	1.561	2.106
Dwelling Unit Features	Area per room (natural log)	29.436	5.112
	Total number of rooms (includes kitchen & bath) (natural log)	71.115	14.915
	Building age (years)	-0 217	1.939
	Stove or refrigerator provided (0,1)	4.066	1.468
	Central heat present (0,1)	8.801	1.854
	Garage or carport provided (0,1)	9.136	3.176
	Dishwasher and/or disposal provided (0,1)	4.517	0.833
	Recent interior painting or papering (0,1)	3.306	1.100
	Average surface and structural quality (4 point scale)	13.106	4 899
	Adequate light and ventilation (0,1)	9 152	3.148
	Central air-conditioning present (0,1)	17.389	2.936
	Large multifamily structure (0,1)	6.522	1.433
Neighborhood Features	Overall neighborhood quality (factor score)	1.443	.842
	Recreational facilities (factor score)	2.195	1.529
	Access to shopping and parking (factor score)	1.467	0.876
	Census tracts with higher priced dwelling units and higher socioeconomic status (factor score)	3.300	1.627
	Owner-occupied, single-family units in census tract (factor score)	1.297	0.826
	Poor quality housing in census tract (factor score)	-2.882	3.323
	Distance from the Central Business District (miles)	-0.179	0.589
	Quality of block face landscaping (4 point scale)	3.710	2.466
CONSTANT		-167 523	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-27

SPANISH HOUSEHOLDS SEMILOG EQUATION: PHOENIX

	R^2 = 0 793 $\bar{R}^2$ = 0 777 F = 51 648 N = 378		
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.056	1.583
	Length of residence (exponential function)	-.207	6.834
	Number of persons per room	.050	2.912
	Number of landlord contacts for maintenance	.015	2.405
Dwelling Unit Features	Area per room (natural log)	.257	5.238
	Total number of rooms (includes kitchen & bath) (natural log)	.688	16 956
	Building age (years)	-.002	2.210
	Stove or refrigerator provided (0,1)	030	1.272
	Central heat present (0,1)	.039	0.977
	Garage or carport provided (0,1)	068	2.813
	Dishwasher and/or disposal provided (0,1)	.007	0.148
	Recent interior painting or papering (0,1)	.039	1.577
	Average surface and structural quality (4 point scale)	.105	4.384
	Adequate light and ventilation (0,1)	046	1.891
	Central air-conditioning present (0,1)	.071	1.437
	Large multifamily structure (0,1)	.071	1.867
	Plumbing present (0,1)	.065	2.306
	Inferior or no heat (0,1)	-.036	1 464
	Presence of adequate ceiling height (0,1)	.058	2.174
Neighborhood Features	Overall neighborhood quality (factor score)	.024	1.733
	Recreational facilities (factor score)	.018	1.644
	Access to shopping and parking (factor score)	030	2.165
	Census tracts with higher priced units and higher socioeconomic status (factor score)	.004	0.207
	Poor quality housing in census tract (factor score)	-.020	2.867
	Distance from the Central Business District (miles)	-.001	0.558
	Quality of block face landscaping (4 point scale)	.031	2.434
CONSTANT		2.076	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-28

WHITE HOUSEHOLDS LINEAR EQUATION: PHOENIX

	R^2 = 0.782	$\bar{R}^2$ = 0.777 F = 155.180 N = 1,065	
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-21.485	7.329
	Length of residence (exponential function)	-23.914	10.299
	Number of persons per room	10.109	5.032
	Number of landlord contacts for maintenance	1.068	2.158
Dwelling Unit Features	Area per room (natural log)	41.467	11.435
	Total number of rooms (includes kitchen & bath) (natural log)	83.992	28.497
	Building age (years)	-0 345	4.760
	Stove or refrigerator provided (0,1)	1 666	0.773
	Central heat present (0,1)	7.051	3.484
	Garage or carport provided (0,1)	4.109	2.789
	Dishwasher and/or disposal provided (0,1)	8.153	4.025
	Recent interior painting or papering (0,1)	0.669	0.406
	Average surface and structural quality (4 point scale)	14.347	7.014
	Adequate light and ventilation (0,1)	6 604	4.586
	Central air-conditioning present (0,1)	6.465	2 862
	Large multifamily structure (0,1)	4.015	1.972
Neighborhood Features	Overall neighborhood quality (factor score)	2.894	3.268
	Recreational facilities (factor score)	2.776	3.582
	Access to shopping and parking (factor score)	1.418	1.638
	Census tracts with higher priced dwelling units and higher socioeconomic status (factor score)	1.708	1.420
	Owner-occupied, single-family units in census tract (factor score)	1.436	1.813
	Poor quality housing in census tract (factor score)	-0.586	0.843
	Distance from the Central Business District (miles)	-0.562	3 044
	Quality of block face landscaping (4 point scale)	1.994	2.402
CONSTANT		-231.464	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-29

WHITE HOUSEHOLDS SEMILOG EQUATION PHOENIX

	R^2 = 0 799 $\bar{R}^2$ = 0 794 F = 158.302 N = 1,065		
VARIABLE DESCRIPTION		COEFFICIENTS	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-.178	7.986
	Length of residence (exponential function)	-.198	11.300
	Number of persons per room	.080	5.307
	Number of landlord contacts for maintenance	014	3.727
Dwelling Unit Features	Area per room (natural log)	.348	12.312
	Total number of rooms (includes kitchen & bath) (natural log)	.682	29.485
	Building age (years)	-.003	6.039
	Stove or refrigerator provided (0,1)	.015	0.931
	Central heat present (0,1)	.033	2.121
	Garage or carport provided (0,1)	.028	2.489
	Dishwasher and/or disposal provided (0,1)	.032	2.115
	Recent interior painting or papering (0,1)	004	0.339
	Average surface and structural quality (4 point scale)	.128	7.608
	Adequate light and ventilation (0,1)	033	3.032
	Central air-conditioning present (0,1)	.051	2.965
	Large multifamily structure (0,1)	.016	1.051
	Plumbing present (0,1)	.043	1 518
	Inferior or no heat (0,1)	-.021	1.336
	Presence of adequate ceiling height (0,1)	-.011	0.560
Neighborhood Features	Overall neighborhood quality (factor score)	.022	3.339
	Recreational facilities (factor score)	.016	2.953
	Access to shopping and parking (factor score)	.016	2.468
	Census tracts with higher priced units and higher socioeconomic status (factor score)	.012	1.322
	Poor quality housing in census tract (factor score)	-.002	0.300
	Distance from the Central Business District (miles)	-.004	3.082
	Quality of block face landscaping (4 point scale)	.014	2.254
CONSTANT		1.796	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-30

MEANS AND STANDARD DEVIATIONS PITTSBURGH
CONTROL HOUSEHOLDS AT ENROLLMENT AND TWO YEARS

VARIABLE DESCRIPTION		ENROLLMENT MEANS	ENROLLMENT STANDARD DEVIATION	TWO YEAR PERIOD MEANS	TWO YEAR PERIOD STANDARD DEVIATION
Tenure Characteristics	Related to landlord (0,1)	0 088	0.283	0.106	0.308
	Length of residence (exponential function)	0.447	0.370	0.546	0.361
	Landlord lives in the building (0,1)	0.110	0.313	0.090	0.287
	Number of persons per room	0.737	0.334	0.683	0.342
	Number of landlord contacts for maintenance	1.276	1.281	0.489	0.916
Dwelling Unit Features	Area per room (natural log)	4.853	0.176	4.866	0.182
	Total number of rooms (includes kitchen and bath) (natural log)	1.684	0.253	1.733	0.297
	Building age (years)	48.705	15.044	50.109	14.955
	Stove and refrigerator provided (0,1)	0 116	0.321	0.118	0 324
	Inferior or no heat (0,1)	0.201	0.401	0.165	0.372
	Garage provided (0,1)	0.072	0.259	0.084	0.278
	Offstreet parking provided (0,1)	0.091	0.288	0.150	0.357
	Overall evaluator rating (4 point scale)	1.868	0.679	1.950	0.557
	Dishwasher and/or disposal provided (0,1)	0.075	0.264	0.084	0.278
	Recent interior painting or papering (0,1)	0.125	0 332	0.068	0.253
	Many high quality features (0,1)	0.060	0.237	0.040	0.197
	Poor wall and ceiling surface (factor score)	-0.120	1.041	-0.356	1.043
	Poor window condition (factor score)	-0.011	0.993	0.043	0.919
	Poor bathroom wall and ceiling surface (factor score)	0.052	1.062	0.015	1.128
	High quality kitchen (0,1)	0.103	0 305	0.081	0.273
	Presence of adequate exits (0,1)	0.934	0.248	0.938	0.242
	Air-conditioning present (0,1)	0.119	0.324	0.115	0.320
	Presence of adequate ceiling height (0,1)	0.900	0.301	0.925	0.263
	Adequate kitchen facilities present (0,1)	0 994	0.079	0.994	0.079
	Large multifamily structure (0,1)	0.144	0.352	0.159	0.366
	Plumbing present (0,1)	3.580	0 824	3 452	1.204
	Presence of private yard (0,1)	0.342	0.475	0 383	0.487
Neighborhood Features	Good recreational facilities and access (factor score)	0.004	0.977	0.123	0.990
	Traffic and litter problems (factor score)	-0.014	0.973	-0.869	0.850
	Problems with crime and public services (factor score)	-0.097	0.962	0 245	0.734
	Census tracts with higher priced units and higher socio-economic status (factor score)	0 0001	0.992	0.156	1.096
	Nonminority census tracts with higher socioeconomic status (factor score)	0.112	0.883	0.062	0.938
	Blue collar workers and nonminority residents in census tract (factor score)	0.128	0.930	0.156	0.897
	High quality block face (0,1)	0.373	0.484	0.498	0.501
Rent	Analytic rent	114 524	34.996	126.143	39 998
	Natural logarithm of analytic rent	4 694	0.302	4.789	0.314

SAMPLE. All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-31

MEANS AND STANDARD DEVIATIONS PHOENIX
CONTROL HOUSEHOLDS AT ENROLLMENT AND TWO YEARS

VARIABLE DESCRIPTION		ENROLLMENT MEANS	ENROLLMENT STANDARD DEVIATION	TWO YEAR PERIOD MEANS	TWO YEAR PERIOD STANDARD DEVIATION
Tenure Characteristics	Related to landlord (0,1)	0.092	0.290	0.075	0.263
	Length of residence (exponential function)	0.376	0.340	0.436	0.356
	Number of persons per room	0.890	0.558	0 753	0.429
	Number of landlord contacts for maintenance	1 244	1.403	0 734	1 247
Dwelling Unit Features	Area per room (natural log)	4.686	0.197	4.669	0.186
	Total number of rooms (includes kitchen & bath) (natural log)	1.600	0.204	1 695	0 239
	Building age (years)	26.260	15.026	25.726	16.400
	Stove or refrigerator provided (0,1)	0.744	0.437	0.718	0.451
	Central heat present (0,1)	0.312	0.464	0.394	0.490
	Garage or carport provided (0,1)	0.316	0.466	0.344	0 476
	Dishwasher and/or disposal provided (0,1)	0 168	0.375	0.203	0 403
	Recent interior painting or papering (0,1)	0.176	0 382	0.100	0.300
	Average surface and structural quality (4 point scale)	2.192	0.691	2.298	0 603
	Adequate light and ventilation (0,1)	0.364	0.482	0.473	0.500
	Central air-conditioning present (0,1)	0.244	0.430	0.299	0.459
	Large multifamily structure (0,1)	0.176	0.382	0.174	0.380
	Plumbing present (0,1)	0.896	0 306	0.942	0.234
	Inferior or no heat (0,1)	0.360	0.481	0.295	0.457
	Presence of adequate ceiling height (0,1)	0.888	0.316	0.867	0.340
Neighborhood Features	Overall neighborhood quality (factor score)	-0.044	1.009	0.344	1.068
	Recreational facilities (factor score)	-0.067	1.028	0.761	0 721
	Access to shopping and parking (factor score)	-0 128	0.998	0 056	0.721
	Census tracts with higher priced units and higher socioeconomic status (factor score)	-0.106	0.977	0.012	1.052
	Owner-occupied single family dwelling units in census tract (factor score)	-0.053	1.001	0.162	0.986
	Poor quality housing in census tract (factor score)	0.058	0.982	-0.025	0.960
	Distance from the Central Business District (miles)	5.159	4.575	5.515	4.419
	Quality of block face landscaping (4 point scale)	1.740	0.831	1.983	0.619
Rent	Analytic rent	126.586	46.470	139.784	51 872
	Natural logarithm of analytic rent	4.770	0.386	4.865	0.402

SAMPLE All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-32

PITTSBURGH LINEAR EQUATION
CONTROL HOUSEHOLDS AT ENROLLMENT AND TWO YEARS

VARIABLE DESCRIPTION		ENROLLMENT COEFFICIENT	ENROLLMENT t-STATISTIC	TWO YEAR PERIOD COEFFICIENT	TWO YEAR PERIOD t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-4.808	1.134	-25.980	5 315
	Length of residence (exponential function)	-18 836	5.575	-14.813	3 618
	Landlord lives in the building (0,1)	-6.939	1 719	-0.307	0.055
	Number of persons per room	2 010	0 527	9.023	2.013
	Number of landlord contacts for maintenance	1.000	1 035	0.031	0
Dwelling Unit Features	Area per room (natural log)	40 125	5 699	15 790	1.804
	Total number of rooms (includes kitchen & bath) (natural log)	62 299	11 704	60 314	10 539
	Building age (years)	-0 282	2.774	-0 330	2.335
	Stove and refrigerator provided (0,1)	28.445	6 305	7 609	1.313
	Inferior or no heat (0,1)	-3 692	1.150	-9 639	2.253
	Garage provided (0,1)	23.526	5.228	16.160	2.915
	Offstreet parking provided (0,1)	4.045	0 916	1 468	0 339
	Overall evaluator rating (4 point scale)	4.176	1.802	7 474	2.324
	Dishwasher and/or disposal provided (0,1)	5 680	1 170	7 752	1 397
	Recent interior painting or papering (0,1)	4.199	1.143	10 743	1 833
	Many high quality features (0,1)	0.863	0.148	13 543	1 673
	Poor wall and ceiling surface (factor score)	-3 075	2.415	-2.247	1 384
	Poor window condition (factor score)	-2.259	1.727	-2.878	1.742
	Poor bathroom wall and ceiling surface (factor score)	-1.705	1.439	-1.642	1.246
	High quality kitchen (0,1)	2.011	0 477	6.122	1 084
	Presence of adequate exits (0,1)	8 271	1 646	6.575	1.013
	Air-conditioning present (0,1)	-1.573	0 377	3.501	0.721
	Presence of adequate ceiling height (0,1)	-0 158	0 045	4.677	0 869
	Adequate kitchen facilities present (0,1)	9 581	0.622	-16 000	0.784
	Large multifamily structure (0,1)	2 402	0.636	5 669	1 195
Neighborhood Features	Good recreational facilities and access (factor score)	3 079	2.398	5 127	3.181
	Traffic and litter problems (factor score)	-0.998	0.658	-1 263	0 607
	Problems with crime and public services (factor score)	-2.231	1.723	-1 038	0 485
	Census tracts with higher priced units and higher socioeconomic status (factor score)	4.091	2 568	6 214	3.672
	Nonminority census tracts with higher socio-economic status (factor score)	3.824	2.337	3 449	1.896
	Blue collar workers and nonminority residents in census tract (factor score)	-3.243	2 466	-2.332	1 390
	High quality block face (0,1)	13.506	4 958	5.843	1.815
	CONSTANT	-201.136		-53.613	
	R^2	0.703		0.670	
	$\bar{R}^2$	0 670		0 633	
	F	21.504		18 327	
	N	324		322	

SAMPLE: All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-33

PITTSBURGH SEMILOG EQUATION:
CONTROL HOUSEHOLDS AT ENROLLMENT AND TWO YEARS

VARIABLE DESCRIPTION		ENROLLMENT		TWO YEAR PERIOD	
		COEFFICIENT	t-STATISTIC	COEFFICIENT	t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-0 046	1 200	-0.222	5 809
	Length of residence (exponential function)	-0 173	5 852	-0.105	3.328
	Landlord lives in the building (0,1)	-0.083	2 360	-0.036	0 900
	Number of persons per room	0.032	0 970	0.096	2.785
	Number of landlord contacts for maintenance	0.007	0 812	-0.0004	0 032
Dwelling Unit Features	Area per room (natural log)	0.344	5 589	0.077	1.138
	Total number of rooms (includes kitchen & bath) (natural log)	0 555	11 527	0.544	12 190
	Building age (years)	-0.002	2 511	-0.002	2 020
	Stove and refrigerator provided (0,1)	0 228	5 797	0.060	1.344
	Inferior or no heat (0,1)	-0.041	1.472	-0 102	3 080
	Garage provided (0,1)	0.166	4.047	0.092	2.156
	Offstreet parking provided (0,1)	0.022	0 564	0.018	0 538
	Overall evaluator rating (4 point scale)	0.048	2.368	0.073	2.927
	Dishwasher and/or disposal provided (0,1)	0.030	0 715	0.017	0 387
	Recent interior painting or papering (0,1)	0 034	1 068	0 100	2.203
	Many high quality features (0,1)	-0.037	0 736	0.049	0 796
	Poor wall and ceiling surface (factor score)	-0.033	2 935	-0 016	1 251
	Poor window condition (factor score)	-0 016	1 441	-0 023	1 796
	Poor bathroom wall and ceiling surface (factor score)	-0.014	1 349	-0.006	0 581
	High quality kitchen (0,1)	-0.010	0 261	0 057	1 318
	Presence of adequate exits (0,1)	0.067	1 522	0.054	1.071
	Air-conditioning present (0,1)	-0 030	0 828	0.040	1.088
	Presence of adequate ceiling height (0,1)	0.008	0 214	0.022	0 546
	Adequate kitchen facilities present (0,1)	0.042	0.310	-0 196	1 250
	Large multifamily structure (0,1)	0.010	0.306	0 023	0.620
	Plumbing present (0,1)	0 008	0 654	-0 008	0.884
	Presence of private yard (0,1)	0.025	1.065	-0 012	0.515
Neighborhood Features	Good recreational facilities and access (factor score)	0 024	2 167	0 041	3.296
	Traffic and litter problems (factor score)	-0 012	0 903	=0 021	1.335
	Problems with crime and public services (factor score)	-0 021	1 875	-0.016	1.002
	Census tracts with higher priced units and higher socioeconomic status (factor score)	0.026	1.904	0.031	2.415
	Nonminority census tracts with higher socio-economic status (factor score)	0 036	2 504	0 035	2 520
	Blue collar workers and nonminority residents in census tract (factor score)	-0.026	2.306	-0 018	1 418
	High quality block face (0,1)	0.119	5 010	0 044	1 792
	CONSTANT	1.931		3.542	
	R^2	0.702		0.687	
	$\overline{R}^2$	0 666		0.650	
	F	19.678		18.458	
	N	319		321	

SAMPLE All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-34

PHOENIX LINEAR EQUATION
CONTROL HOUSEHOLDS AT ENROLLMENT AND TWO YEARS

VARIABLE DESCRIPTION		ENROLLMENT COEFFICIENT	ENROLLMENT t-STATISTIC	TWO YEAR PERIOD COEFFICIENT	TWO YEAR PERIOD t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-8.740	1.962	-10.046	1.317
	Length of residence (exponential function)	-16.589	4.060	-15.078	2.257
	Number of persons per room	4.710	1.685	12.874	2.232
	Number of landlord contacts for maintenance	0 145	0.158	2.492	1.502
Dwelling Unit Features	Area per room (natural log)	30.980	4.222	32.946	2.748
	Total number of rooms (includes kitchen & bath) (natural log)	97 865	13.213	85 614	8.811
	Building age (years)	-0.533	3.798	-0.135	0.639
	Stove or refrigerator provided (0,1)	4.383	1.257	6.620	1.247
	Central heat present (0,1)	-3.294	0.684	-9 997	1.523
	Garage or carport provided (0,1)	5.276	1.660	11.772	2 508
	Dishwasher and/or disposal provided (0,1)	14.675	3.354	9.004	1.300
	Recent interior painting or papering (0,1)	0.016	0	12.411	1.857
	Average surface and structural quality (4 point scale)	12.752	3.787	16.224	2.700
	Adequate light and ventilation (0,1)	0.750	0.251	8.850	1.954
	Central air-conditioning present (0,1)	11.293	2.160	25.388	3.380
	Large multifamily structure (0,1)	6.170	1.570	-10.237	1.542
Neighborhood Features	Overall neighborhood quality (factor score)	2.674	1.489	1.417	0.462
	Recreational facilities (factor score)	3 517	2.321	3.352	1.063
	Access to shopping and parking (factor score)	1.792	1.069	6.818	1.919
	Census tracts with higher priced units and higher socioeconomic status (factor score)	6.655	2.870	2.030	0.558
	Owner-occupied single family dwelling units in census tract (factor score)	3.626	2.081	0.753	0.264
	Distance from the Central Business District (miles)	-0.552	1.705	-0.491	0.902
	Quality of block face landscaping (4 point scale)	3.023	1.796	9.606	2.435
	CONSTANT	-197 839		-234.955	
	R^2	0.840		0.702	
	$\bar{R}^2$	0.824		0.670	
	F	51.729		22.232	
	N	250		241	

SAMPLE: All enrolled Control households, active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table II-35

PHOENIX SEMILOG EQUATION
CONTROL HOUSEHOLDS AT ENROLLMENT AND TWO YEARS

VARIABLE DESCRIPTION		ENROLLMENT COEFFICIENT	ENROLLMENT t-STATISTIC	TWO-YEAR PERIOD COEFFICIENT	TWO-YEAR PERIOD t-STATISTIC
Tenure Characteristics	Related to landlord (0,1)	-0.077	1.993	-0.006	0.095
	Length of residence (exponential function)	-0.158	4.550	-0.126	2.260
	Number of persons per room	0 042	1 771	0.068	1.429
	Number of landlord contacts for maintenance	0.012	1.507	0.030	2.193
Dwelling Unit Features	Area per room (natural log)	0.299	4.709	0.244	2 383
	Total number of rooms (includes kitchen and bath) (natural log)	0. 0.836	13.000	0.640	7.947
	Building age (years)	-0.004	3.706	-0.001	0.691
	Stove or refrigerator provided (0,1)	0.033	1 099	0 053	1.204
	Central heat present (0,1)	-0.040	0.958	-0.090	1.625
	Garage or carport provided (0,1)	0.062	2.310	0.105	2.715
	Dishwasher and/or disposal provided (0,1)	0.067	1.815	0.058	1.020
	Recent interior painting or papering (0,1)	0.008	0.263	0.058	1.045
	Average surface and structural quality (4 point scale)	0.102	3.208	0.089	1.683
	Adequate light and ventilation (0,1)	-0.030	1.210	0.038	1.016
	Central air-conditioning present (0,1)	0.109	2.472	0.170	2.751
	Large multifamily structure (0,1)	0.043	1.276	-0.047	0 844
	Plumbing present (0,1)	0.104	2.550	0.036	0.425
	Inferior or no heat (0,1)	-0.023	0.733	-0.037	0.768
	Presence of adequate ceiling height (0,1)	-0.008	0.210	0.113	2.138
Neighborhood Features	Overall neighborhood quality (factor score)	0.022	1.459	0.037	1.408
	Recreational facilities (factor score)	0.022	1.667	0 014	0.562
	Access to shopping and parking (factor score)	0.023	1.548	0.034	1.089
	Census tracts with higher priced units and higher socioeconomic status (factor score)	0.035	1.719	0.014	0 456
	Owner-occupied single family dwelling units in census tract (factor score)	0.011	0.748	0.017	0.698
	Poor quality housing in census tract (factor score)	-0.012	0.915	-0.045	2.219
	Distance from the Central Business District (miles)	-0.003	1.258	-0.004	0.930
	Quality of block face landscaping (4 point scale)	0.018	1.246	0.041	1.251
	CONSTANT	1.790		2.121	
	R^2	0.839		0.672	
	$\bar{R}^2$	0.820		0.631	
	F	42.910		16.181	
	N	250		241	

SAMPLE: All enrolled Control households active at two years, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, and those living in their own home or subsidized housing

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

APPENDIX III

ANALYSIS OF UTILITY AND FURNISHINGS ADJUSTMENT ERRORS

As discussed in Section 3.4 of Chapter 3, if rent did not include electricity, gas, heating fuel, water, or garbage collection, the contract rent was adjusted using site-specific utility tables (these are listed in Table VI-1 in Appendix VI, Master Variable List). These tables are based on the number of rooms defined as useable living space (excluding bathrooms, half-rooms, unfinished basements or attics). If the dwelling unit was furnished, contract rent adjusted for utilities was reduced by 11.5 percent to exclude the cost of furnishings.

Estimations of errors in the adjustments were based on the linear model

$$R = X\beta + Z(\gamma_A - \gamma) + \varepsilon \tag{1}$$

where

X = housing, neighborhood, and tenure attributes

Z = utilities not included in rent interacted with unit size, furnishings included in rent interacted with unit size

γ = true utilities and furnishings costs

γ_A = adjustments for utilities (from tables) or for furnishings (11.5 percent reduction of adjusted rent).

The data for utility and furnishing adjustments are very collinear and the incidence of adjustment is strongly related to the number of rooms in the dwelling unit. (See Figures 3-6 and 3-7, Chapter 3.) For example, although the three major utilities--gas, heat, and electricity--could be included or excluded from the contract rent in a variety of ways, Tables III-1 and III-2 indicate that the majority of units in each unit size had either all or none of these utilities included in contract rent, and the tendency to have these utilities included in rent decreased as unit size increased.

Because of the collinearity of these data, there are potentially six types of adjustments: gas, heat, electricity, water, garbage collection, and furnishings. In addition, for all utilities but garbage collection the amount of the adjustment varied by unit size. Among the various specifications were:

Table III-1

PERCENT OF PITTSBURGH UNITS WITH UTILITIES INCLUDED IN RENT, BY UNIT SIZE AND TYPE OF UTILITY

Number of Rooms	(N)	No Utilities Included	Gas Included	Heat Included	Electricity Included	Gas, Heat Included	Electricity, Heat Included	Gas, Electricity Included	Everything Included	Everything Included or Nothing Included
1, 2	69	17.4%	-	1.4	-	7 2	1 4	-	72.5	89 9
3	267	40.4	1.9	8 2	4	10 9	-	-	38 2	78 6
4	588	67.9	1.7	7 1	7	9 7	.5	-	12 4	80 3
5	375	70.4	2 4	10.4	-	6 4	1 1	-	9.3	79 7
6	223	86 5	1.8	3.1	-	1.8	-	-	6 7	93.2
7+	92	88.0	2.2	7.6	-	1.1	-	-	1 1	89 1

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent.

DATA SOURCES Initial Household Report Form, Housing Evaluation Form.

Table III-2

PERCENT OF PHOENIX UNITS WITH UTILITIES INCLUDED IN RENT, BY UNIT SIZE AND TYPE OF UTILITY

Number of Rooms	(N)	No Utilities Included	Gas Included	Heat Included	Electricity Included	Gas, Heat Included	Electricity, Heat Included	Gas, Electricity Included	Everything Included	Everything Included or Nothing Included
1,2	84	28.6%	-	4.8	-	-	-	-	66.7	95.3
3	385	53.5	6.5	4.7	.5	2.9	.3	-	31.7	85.2
4	677	66.9	7.5	5.6	-	3.8	.1	-	16.0	82.9
5	265	68.7	4.9	8.7	.4	3.4	-	-	14.0	82.7
6	102	83.3	4.9	3.9	-	2.0	-	-	5.9	89.2
7+	30	93.3	-	3.3	-	-	-	-	3.3	96.6

SAMPLE. All enrolled households excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES Initial Household Report Form, Housing Evaluation Form.

(0,1) dummy to indicate the presence ("1") of any adjustment to utilities or furnishings;

(0,1) dummies to indicate the presence ("1") of each of the individual utilities or furnishings in contract rent;

(0,1) dummies to indicate the presence ("1") of combinations of individual utilities in contract rent (e.g., gas, heat, and electricity);

(0,1) dummies to indicate the presence ("1") of combinations of individual utilities in contract rent, interacted by unit size (e.g., gas, heat, and electricity for four-room units; all utilities for five-room units);

interaction of number of rooms with individual or combinations of utility dummies (e.g., number of rooms x gas, heat, and electricity);

interaction of logarithm of number of rooms with individual or combined utility dummies (e.g., ln(number of rooms) x (gas, heat, and electricity)).

It was impossible to obtain specific error estimates for each of the individual utility variables due to their collinearity. Variables indicating the presence of adjustments for garbage collection were eliminated from all equations after initial attempts indicated that the error in this adjustment could not be estimated (garbage collection is almost always included, though inclusion does fall somewhat with distance from the Central Business District). In addition, many of the combinations of utility variables were insignificant in the equations. Ultimately, two sets of variables were selected. The first consists of a dummy variable indicating that any adjustment was made; that is, an adjustment to gas, heat, electricity, water, or furnishings. Also, an interaction term of this dummy variable with the number of rooms is included. The second set of variables consists of a variable combining the three main utility variables (gas, heat, and electricity) to indicate that presence of adjustments for all three variables. Water is estimated separately. The variables for water and for the three main utilities are both interacted with number of rooms. Furnishings are estimated by a single dummy variable.[1]

[1]In Pittsburgh, since only 4 percent of all units were furnished, small sample sizes prevented any interaction term of furnishings error with number of rooms. Although almost one-third of all Phoenix units were furnished, the vast majority of these were small units, therefore preventing an accurate specification of an interaction term with number of rooms.

To further assess the accuracy of estimation, the equations were estimated on five separate populations:

1. General - (Pittsburgh = 1,570; Phoenix = 1,524). This sample consisted of all households not in their own home or subsidized housing, that did not receive a reduced rent because they worked for their landlord, that did not move between the Baseline Interview and enrollment and that had extreme values for residuals in the final hedonic equation. The first three selection criteria eliminated households that had extreme or missing values for rent. The fourth selection eliminated a small number of households for whom the model predicted poorly.

2. Consistent Utility Payments - (Pittsburgh = 1,295; Phoenix = 1,293). This sample selected from the general sample all households who paid for gas, heat, and electricity, and those who did not pay for any of the three. As indicated in Tables III-1 and III-2, this included the majority of units. The hypothesis was that the model would predict errors in utility adjustments most accurately if it was estimated on units that either had no additional payments for the three utilities (and therefore no adjustments made to these utilities), or had all three utilities but paid for them separately (and therefore adjustments made for all three utilities).

3. Dwelling Unit Size - (Pittsburgh = 1,432; Phoenix = 1,417). This sample selected from the general sample all dwelling units of three, four, five, or six rooms. The hypothesis was that the model would predict poorly for very small or very large units, because the samples for these unit sizes are small. (A sample of three-, four-, and five-room units was also tested, but no significant differences were indicated.)

4. Stratification by Dwelling Unit Size - This selected from the general sample individual samples of one- and two-, three-, four-, five-, six-room units to produce samples containing all households living in the same size dwelling unit.

5. <u>Unfurnished Units</u> (Pittsburgh = 1,513; Phoenix = 1,037). This sample selected from the general sample all unfurnished dwelling units. The hypothesis was that the model would predict errors in utility adjustments most accurately in the absence of furnishings adjustments.

A number of the more restricted samples provided little improvement in estimation over the general sample. Most of the estimates indicated similar error patterns, although the exact amounts of these estimates differed. Two estimates were finally selected; an error estimate indicating the average adjustment error for all households in the general sample (sample 1) that had any adjustment and a more specific error estimate indicating the adjustment error for water, for furnishings, and for the combination of gas, heat, and electricity for the restricted sample of households with consistent utility payments (sample 2). Also, the average adjustment error was estimated for households in unfurnished units in Phoenix (sample 5).

The results of the final equations are listed in Tables III-3 to III-7. Note that the coefficients of the adjustment/room interactions are small relative to the coefficients for total rooms. Most of the coefficients are significant and the adjusted $\bar{R}^2$s approximate those for the general linear model (see Tables 3-3 and 3-5, Chapter 3).

For the restricted sample, the individual dummy variables for utilities and furnishings are equal to "1" when the utility or furnishings do not require additional payment beyond contract rent. (Since, analytically, the coefficient for these variables is to be interpreted as the estimated adjustment error, the signs of the coefficients for the utility variables are reversed in the text in Chapter 3 to indicate the direction of error made in adjusting contract rent when it does not include utilities.)

The coefficients for the variables interacting adjustments with number of rooms are interpreted as adjustment errors per unit size. For example, the adjustment error for water for a five-room unit would be estimated by multiplying the coefficient of the water/room interaction by the unit size, 5. The coefficients of the dummy variables indicating presence of adjustments are interpreted as base adjustment errors. These estimates are added to the error estimates from the adjustment/room interactions to produce a

Table III-3

ESTIMATION OF GENERAL ADJUSTMENT ERROR
(Dependent variable: analytic rent)
PITTSBURGH

$R^2 = 0.664$ | $\bar{R}^2 = 0.657$ | $F = 89.371$ | $N = 1,570$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-13 031	6.604
	Length of residence (exponential function)	-14.727	10.886
	Landlord lives in the building (0,1)	-4.873	2.859
	Number of persons per room	7.716	5.021
	Number of landlord contacts for maintenance	1.082	2.974
Dwelling Unit Features	Area per room (natural log)	18.518	6.296
	Total number of rooms (includes kitchen & bath) (natural log)	48 008	11.566
	Building age (years)	-0.231	5 185
	Stove and refrigerator provided (0,1)	14.067	7 357
	Inferior or no heat (0,1)	-6.682	5 023
	Garage provided (0,1)	14 772	7 261
	Offstreet parking provided (0,1)	2.412	1.335
	Overall evaluator rating (4 point scale)	5 532	5.656
	Dishwasher and/or disposal provided (0,1)	10 337	4.624
	Recent interior painting or papering (0,1)	6 995	4 239
	Many high quality features (0,1)	7.667	2.835
	Poor wall and ceiling surface (factor score)	-1.750	3 321
	Poor window condition (factor score)	-2.117	3.917
	Poor bathroom wall and ceiling surface (factor score)	-1.492	3 089
	High quality kitchen (0,1)	5.922	3.081
	Presence of adequate exits (0,1)	3 998	2.094
	Air-conditioning present (0,1)	2.920	1 780
	Presence of adequate ceiling height (0,1)	2.308	1 343
	Adequate kitchen facilities present (0,1)	7 671	1.370
	Large multifamily structure (0,1)	3 572	2 162
Neighborhood Features	Good recreational facilities and access (factor score)	2 510	4.758
	Traffic and litter problems (factor score)	-1 489	2.410
	Problems with crime and public services (factor score)	-1 730	3 041
	Census tracts with higher priced units and higher socioecnomic status (factor score)	3 787	6 094
	Nonminority census tracts with higher socioeconomic status (factor score)	3.819	6 063
	Blue collar workers and nonminority residents in census tracts (factor score)	-2.652	5.366
	High quality block face (0,1)	5 125	4.506
Utility and Furnishings Adjustment	Adjustment for gas, heat, electricity, water, or furnishings (0,1)	-4 623	1.217
	Adjustment for gas, heat, electricity, water, or furnishings x number of rooms	2 635	2 981
CONSTANT		-82.59745	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent

DATA SOURCES. Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population

a. A t-statistic $\geq$ 1 0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0 125 level of confidence for a one-tailed test

Table III-4

ESTIMATION OF GENERAL ADJUSTMENT ERROR
(Dependent variable. analytic rent)
PHOENIX

$R^2 = 0.799$ $\bar{R}^2 = 0.795$ $F = 228.327$ $N = 1,524$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-14.989	6.393
	Length of residence (exponential function)	-22.099	11.991
	Number of persons per room	8.290	6.481
	Number of landlord contacts for maintenance	1.179	2.956
Dwelling Unit Features	Area per room (natural log)	32.445	10.806
	Total number of rooms (includes kitchen & bath) (natural log)	50.265	11.353
	Building age (years)	-0 288	5.039
	Stove or refrigerator provided (0,1)	5.301	3.329
	Central heat present (0,1)	7.142	3.993
	Garage or carport provided (0,1)	4 107	3.249
	Dishwasher and/or disposal provided (0,1)	8.162	4.375
	Recent interior painting or papering (0,1)	2.983	2.158
	Average surface and structural quality (4 point scale)	14.156	9.318
	Adequate light and ventilation (0,1)	6.413	5.185
	Central air-conditioning present (0,1)	9.441	4.661
	Large multifamily structure (0,1)	4.488	2.484
Neighborhood Features	Overall neighborhood quality (factor score)	2.779	3.842
	Recreational facilities (factor score)	2.420	3.733
	Access to shopping and parking (factor score)	1.265	1.704
	Census tracts with higher priced units and higher socioeconomic status (factor score)	3.593	3.787
	Owner-occupied single family dwelling units in census tracts (factor score)	1.201	1.757
	Poor quality housing in census tracts (factor score)	-3.119	4.711
	Distance from the Central Business District (miles)	-0.488	3.312
	Quality of block face landscaping (4 point scale)	2.724	3.959
Utility and Furnishings Adjustments	Adjustment for gas, heat, electricity, water or furnishings (0,1)	-33.570	7.450
	Adjustment for gas, heat, electricity, water or furnishings x number of rooms	8.560	7.809
CONSTANT		-144.950	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic ≥ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table III-5

ESTIMATION OF GENERAL ADJUSTMENT ERROR
(Dependent variable analytic rent)
PHOENIX UNFURNISHED UNITS

$R^2 = 0.806$ $\bar{R}^2 = 0.800$ $F = 161.016$ $N = 1{,}037$

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-15.840	5 727
	Length of residence (exponential function)	-23.832	10 532
	Number of persons per room	7.590	4.860
	Number of landlord contacts for maintenance	1.156	2.418
Dwelling Unit Features	Area per room (natural log)	32.278	7.214
	Total number of rooms (includes kitchen & bath) (natural log)	55.344	10.735
	Building age (years)	-0.202	2.719
	Stove or refrigerator provided (0,1)	8.237	4.584
	Central heat present (0,1)	8.696	3.796
	Garage or carport provided (0,1)	3.827	2.528
	Dishwasher and/or disposal provided (0,1)	6.970	3.159
	Recent interior painting or papering (0,1)	2.149	1 230
	Average surface and structural quality (4 point scale)	15.539	8.292
	Adequate light and ventilation (0,1)	6.600	4.189
	Central air-conditioning present (0,1)	7.304	2.885
	Large multifamily structure (0,1)	6.158	2.384
Neighborhood Features	Overall neighborhood quality (factor score)	2.795	3.074
	Recreational facilities (factor score)	2.639	3.269
	Access to shopping and parking (factor score)	1.173	1.258
	Census tracts with higher priced units and higher socioeconomic status (factor score)	3.805	3.152
	Owner-occupied single family dwelling units in census tracts (factor score)	0.530	0.607
	Poor quality housing in census tracts (factor score)	2.518	3.049
	Distance from the Central Business District (miles)	-0.430	2.317
	Quality of block face landscaping (4 point scale)	3.648	4.158
Utility and Furnishings Adjustments	Adjustment for gas, heat, electricity, water or furnishings (0,1)	-25.371	4.919
	Adjustment for gas, heat, electricity, water or furnishings x number of rooms	7 156	5.733
CONSTANT		-160.450	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, those reporting work for landlord in lieu of rent, and those living in a unit with furnishings included in contract rent.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table III-6

ESTIMATION OF SPECIFIC ADJUSTMENT ERRORS
(Dependent variable. analytic rent)
PITTSBURGH

	$R^2 = 0.657$	$\bar{R}^2 = 0.647$	$F = 66.954$	N = 1,295	

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-13 218	6 400
	Length of residence (exponential function)	-15.954	10.962
	Landlord lives in the building (0,1)	-1.978	1 024
	Number of persons per room	7.191	4.383
	Number of landlord contacts for maintenance	0 900	2.276
Dwelling Unit Features	Area per room (natural log)	16.725	5 316
	Total number of rooms (includes kitchen & bath) (natural log)	60.889	23.665
	Building age (years)	-0.197	3 998
	Stove and refrigerator provided (0,1)	16.672	6.413
	Inferior or no heat (0,1)	-8.493	6.009
	Garage provided (0,1)	15 594	7.336
	Offstreet parking provided (0,1)	-0.167	0.084
	Overall evaluator rating (4 point scale)	5.091	4.834
	Dishwasher and/or disposal provided (0,1)	8 033	3.286
	Recent interior painting or papering (0,1)	6 232	3.426
	Many high quality features (0,1)	5 972	1.992
	Poor wall and ceiling surface (factor score)	-1 223	2.151
	Poor window condition (factor score)	-1 827	3.169
	Poor bathroom wall and ceiling surface (factor score)	-1.615	3 179
	High quality kitchen (0,1)	4 401	2.087
	Presence of adequate exits (0,1)	0 430	0.187
	Air-conditioning present (0,1)	0 744	0.392
	Presence of adequate ceiling height (0,1)	3.657	2.032
	Adequate kitchen facilities present (0,1)	7 966	1 385
	Large multifamily structure (0,1)	2.556	1 208
Neighborhood Features	Good recreational facilities and access (factor score)	3.092	5.351
	Traffic and litter problems (factor score)	-1.373	2.044
	Problems with crime and public services (factor score)	-0.820	1.316
	Census tracts with higher priced units and higher socioeconomic status (factor score)	3.521	5.007
	Nonminority census tracts with higher socio-economic status (factor score)	3 545	5.204
	Blue collar workers and nonminority residents in census tracts (factor score)	-2.392	4 474
	High quality block face (0,1)	4.311	3 483
Utility and Furnishings Adjustments	No additional payment for gas, heat, and electricity (0,1)	-5.358	1.033
	No additional payment for gas, heat, and electricity x number of rooms	-0 418	0 330
	No additional payment for water included x number of rooms	-0 819	3.324
	Furnishings included in rent (0,1)	-4.832	1 397
CONSTANT		-81 266	

SAMPLE All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic ≥ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table III-7

ESTIMATION OF SPECIFIC ADJUSTMENT ERRORS
(Dependent variable analytic rent)
PHOENIX

R^2 = 0 792 $\bar{R}^2$ = 0 788 F = 172 307 N = 1,293

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-13.555	5.271
	Length of residence (exponential function)	-24.461	11.999
	Number of persons per room	8.281	6.055
	Number of landlord contacts for maintenance	1.026	2.335
Dwelling Unit Features	Area per room (natural log)	33.332	9.911
	Total number of rooms (includes kitchen & bath) (natural log)	79 578	25.193
	Building age (years)	-0.248	3.968
	Stove or refrigerator provided (0,1)	6.098	3.506
	Central heat present (0,1)	7.885	3.962
	Garage or carport provided (0,1)	4.133	2.906
	Dishwasher and/or disposal provided (0,1)	7.540	3.387
	Recent interior painting or papering (0,1)	2 876	1 848
	Average surface and structural quality (4 point scale)	14.313	8.698
	Adequate light and ventilation (0,1)	5.948	4.258
	Central air-conditioning present (0,1)	8.765	3.739
	Large multifamily structure (0,1)	4 640	2.173
Neighborhood Features	Overall neighborhood quality (factor score)	2.715	3.367
	Recreational facilities (factor score)	2.879	3.966
	Access to shopping and parking (factor score)	0.580	0.698
	Census tracts with higher priced units and higher socioeconomic status (factor score)	4.325	4.107
	Owner-occupied single family dwelling units in census tracts (factor score)	0.846	1.089
	Poor quality housing in census tracts (factor score)	-3.088	4.264
	Distance from the Central Business District (miles)	-0.393	2.340
	Quality of block face landscaping (4 point scale)	2 335	3.078
Utility and Furnishings Adjustments	No additional payment for gas, heat, and electricity (0,1)	12.016	2.169
	No additional payment for gas, heat, and electricity x number of rooms	-3.213	2.172
	No additional payment for water included x number of rooms	-0 271	0.750
	Furnishings included in rent (0,1)	-8.686	5.732
CONSTANT		-191.981	

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

total estimated error of the adjustments for each unit size.[1]

Although the results presented in Chapter 3 used analytic rent rather than monthly contract rent as the dependent variable in order to obtain direct estimates of error, utility costs can also be estimated for the restricted sample by regressing unadjusted monthly contract rent on unit characteristics and utility/furnishings variables. Again, using the terms developed in the model presented in Section 3.4, the estimating equation for this specification is

$$R_c = X\beta + Z\lambda + \varepsilon$$

where

R_c = monthly contract rent

Z = utilities, furnishings included in contract rent, interacted with unit size

λ = utility, furnishings cost.

The results are presented in Tables III-8 and III-9. The difference between these cost estimates and the utility table entries are not necessarily the same as the error estimates derived from analytic rent.[2] The difference between the estimated costs of gas, heat, electricity, and water and those costs used in the utility tables is illustrated in Figures III-1 and III-2. In Pittsburgh, the difference is almost constant for all unit sizes; in Phoenix, the estimated costs are similar to the utility tables for all unit sizes except very large units. In fact, the error measured by differences between the cost estimates and the utility tables is similar to the estimated errors, as shown in Table III-10.

---

[1]The predicted error in furnishings adjustment is difficult to evaluate; the furnishing adjustment reduces the sum of contract rent plus utilities by a constant percentage and the dollar value of this percentage is not available as a unique variable. The adjustments are made during the derivation of analytic rent, and only the final analytic rent value exists as a unique variable.

[2]The two would necessarily provide identical estimates only if the structure of utility/furnishings variables exactly mirrors the structure of the adjustments.

Table III-8

ESTIMATION OF THE COST OF UTILITIES
(Dependent variable. unadjusted monthly contract rent)
PITTSBURGH

	R^2 = 0 626	$\bar{R}^2$ = 0 616	F = 58 586	N = 1,295

VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-12.828	6 244
	Length of residence (exponential function)	-16.593	11 460
	Landlord lives in the building (0,1)	-2 146	1.117
	Number of persons per room	7 285	4.463
	Number of landlord contacts for maintenance	0.851	2.164
Dwelling Unit Features	Area per room (natural log)	13 548	4.329
	Total number of rooms (includes kitchen & bath) (natural log)	37 240	14.548
	Building age (years)	-0.181	3 686
	Stove and refrigerator provided (0,1)	16.254	6.284
	Inferior or no heat (0,1)	-8.072	5.740
	Garage provided (0,1)	14 912	7 052
	Offstreet parking provided (0,1)	0.086	0 045
	Overall evaluator rating (4 point scale)	5 349	5.105
	Dishwasher and/or disposal provided (0,1)	7 407	3.045
	Recent interior painting or papering (0,1)	6 235	3.444
	Many high quality features (0,1)	5.510	1.847
	Poor wall and ceiling surface (factor score)	-1.293	2 286
	Poor window condition (factor score)	-1.873	3.265
	Poor bathroom wall and ceiling surface (factor score)	-1.422	2.815
	High quality kitchen (0,1)	4.600	2.192
	Presence of adequate exits (0,1)	0.125	0.055
	Air-conditioning present (0,1)	0.993	0 526
	Presence of adequate ceiling height (0,1)	3.052	1 705
	Adequate kitchen facilities present (0,1)	9.476	1.656
	Large multifamily structure (0,1)	2 045	0.971
Neighborhood Features	Good recreational facilities and access (factor score)	3 216	5.594
	Traffic and litter problems (factor score)	-1.393	2.085
	Problems with crime and public services (factor score)	-0.757	1.220
	Census tracts with higher priced units and higher socioeconomic status (factor score)	3 341	4 842
	Nonminority census tracts with higher socio-economic status (factor score)	3.867	5 706
	Blue collar workers and nonminority residents in census tracts (factor score)	-2.504	4 708
	High quality block face (0,1)	4.322	3 509
Utility and Furnishings Adjustments	No additional payment for gas, heat, and electricity (0,1)	3 934	0 763
	No additional payment for gas, heat, and electricity x number of rooms	3 687	2.931
	No additional payment for water included x number of rooms	0 559	2 279
	Furnishings included in rent (0,1)	4.512	1.322
CONSTANT		-61 497	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Table III-9

ESTIMATION OF THE COST OF UTILITIES
(Dependent variable unadjusted monthly contract rent)
PHOENIX

R^2 = 0 785 $\bar{R}^2$ = 0 780 F = 165 083 N = 1,293

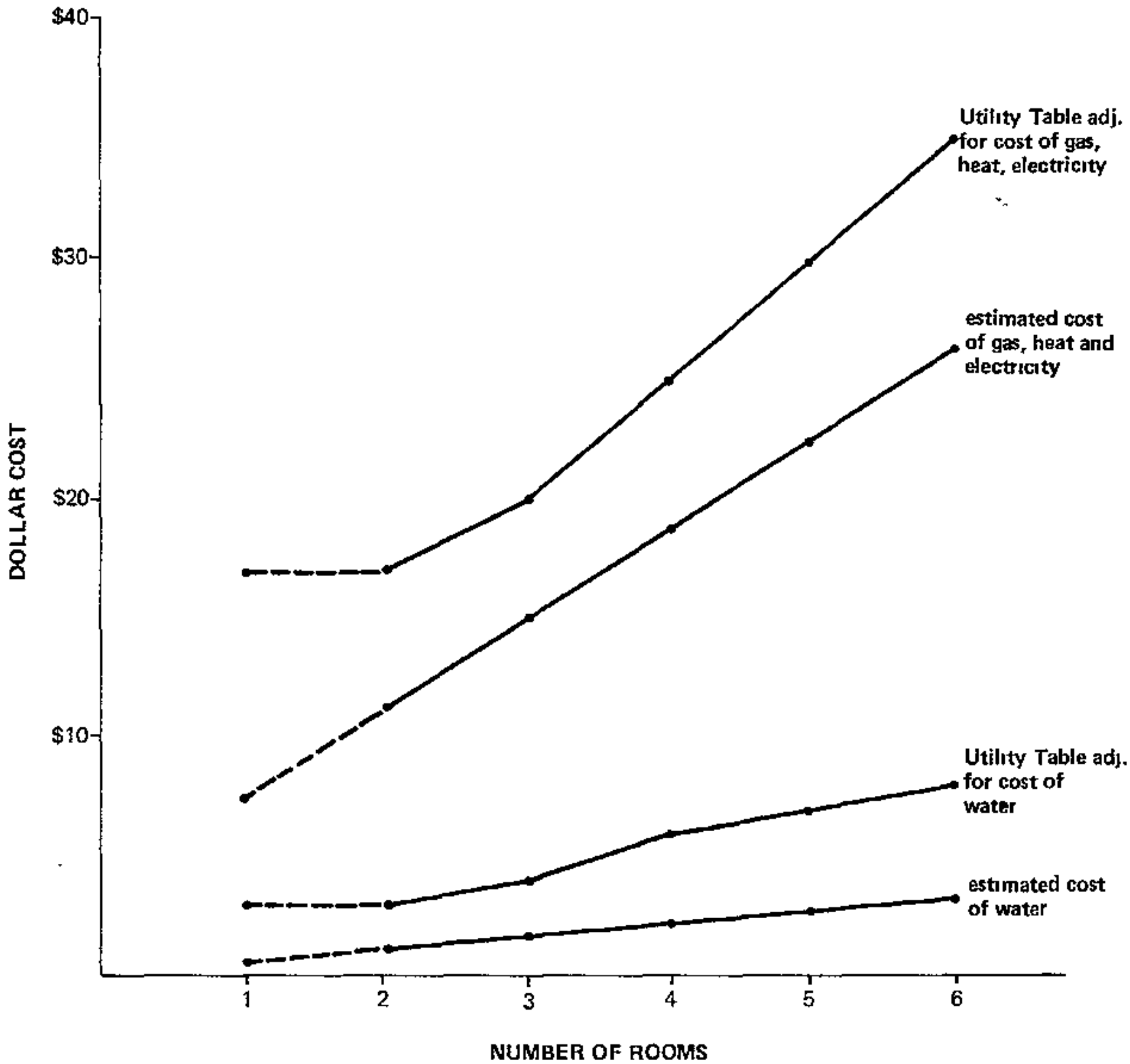
VARIABLE DESCRIPTION		COEFFICIENT	t-STATISTIC[a]
Tenure Characteristics	Related to landlord (0,1)	-14.926	5.731
	Length of residence (exponential function)	-24.868	12.046
	Number of persons per room	9 227	6.662
	Number of landlord contacts for maintenance	1.112	2.500
Dwelling Unit Features	Area per room (natural log)	31.499	9.249
	Total number of rooms (includes kitchen & bath) (natural log)	51.179	15.999
	Building age (years)	-0.291	4.599
	Stove or refrigerator provided (0,1)	6.490	3 684
	Central heat present (0,1)	7.451	3.697
	Garage or carport provided (0,1)	4.158	2.888
	Dishwasher and/or disposal provided (0,1)	7 606	3.374
	Recent interior painting or papering (0,1)	3.229	2.050
	Average surface and structural quality (4 point scale)	15.205	9 124
	Adequate light and ventilation (0,1)	5.564	3 934
	Central air-conditioning present (0,1)	10.281	4.331
	Large multifamily structure (0,1)	5.402	2.498
Neighborhood Features	Overall neighborhood quality (factor score)	3 032	3.713
	Recreational facilities (factor score)	2.906	3 953
	Access to shopping and parking (factor score)	0.997	1 186
	Census tracts with higher priced units and higher socioeconomic status (factor score)	4.011	3.761
	Owner-occupied single family dwelling units in census tracts (factor score)	0.772	0.983
	Poor quality housing in census tracts (factor score)	-3.524	4.806
	Distance from the Central Business District (miles)	-0.434	2.553
	Quality of block face landscaping (4 point scale)	1.972	2.568
Utility and Furnishings Adjustments	No additional payment for gas, heat, and electricity (0,1)	7.836	1.397
	No additional payment for gas, heat, and electricity x number of rooms	4.750	3.171
	No additional payment for water included x number of rooms	1.256	3 428
	Furnishings included in rent (0,1)	7.683	5 007
CONSTANT		-173.654	

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

a. A t-statistic $\geq$ 1.0 indicates significance at the 0.25 level of confidence for a two-tailed test and 0.125 level of confidence for a one-tailed test.

Figure III-1
COMPARISON OF ESTIMATED UTILITY COSTS TO UTILITY TABLES
PITTSBURGH

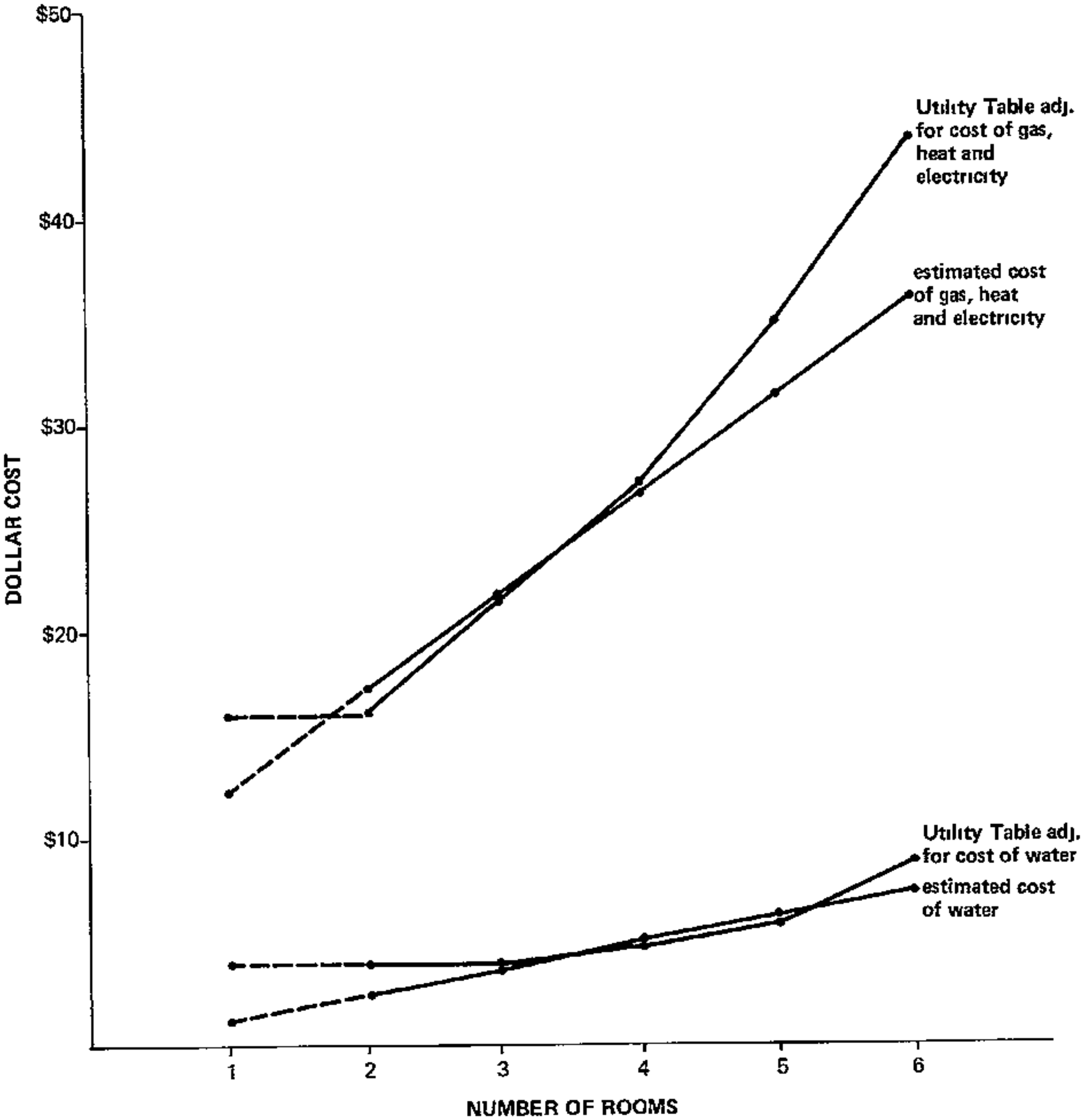


SAMPLE All enrolled households who have no additional payments for the three major utilities — gas, heat, electricity — or who pay extra for all three utilities, excluding those that moved between the Baseline Interviews and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES Initial Household Report Form, Housing Evaluation Form

NOTE Dotted line indicates that there were less than 15 one-room dwelling units in the sample

Figure III-2
COMPARISON OF ESTIMATED COSTS TO UTILITY TABLES
PHOENIX



SAMPLE: All enrolled households who have no additional payments for the three major utilities — gas, heat, electricity — or who pay extra for all three utilities, excluding those that moved between the Baseline Interviews and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for the landlord in lieu of rent

DATA SOURCES Initial Household Report Form, Housing Evaluation Form

NOTE Dotted line indicates that there were less than 15 one-room dwelling units in the sample

Table III-10

COMPARISON OF ESTIMATED UTILITY COSTS
TO UTILITY TABLES AND ESTIMATED ADJUSTMENT ERROR

GAS, HEAT, AND ELECTRICITY

DWELLING UNIT SIZE (Rooms)	2	3	4	5	6
PITTSBURGH					
Estimated cost (predicted with monthly contract rent)	11.31	15.00	18.69	22.38	26.07
Error derived from difference between utility table and estimated cost	5.69	5.00	6.31	7.62	8.93
Estimated error (predicted with analytic rent)	6.20	6.62	7.04	7.46	7.88
PHOENIX					
Estimated cost (predicted with monthly contract rent)	17.34	22.09	26.84	31.59	36.34
Error derived from difference between estimated cost and utility table	-1.34	-0.09	0.16	3.41	7.66
Estimated error (predicted with analytic rent)	-5.60	-2.39	0.82	4.03	7.24

WATER

DWELLING UNIT SIZE (rooms)	2	3	4	5	6
PITTSBURGH					
Estimated cost (predicted with monthly contract rent)	1.12	1.68	2.24	2.80	3.36
Error derived from difference between estimated cost and utility table	1.88	2.32	3.76	4.20	4.64
Estimated error (predicted with analytic rent)	1.64	2.46	3.28	4.10	4.92
PHOENIX					
Estimated cost (predicted with monthly contract rent)	2.52	3.78	5.04	6.30	7.56
Error derived from difference between utility table and estimated cost	1.48	0.22	-0.04	-0.30	1.44
Estimated error (predicted with analytic rent)	0.54	0.81	1.08	1.35	1.62

SAMPLE: All enrolled households who have no additional payments for the three major utilities--gas, heat, electricity--or who pay extra for all three utilities, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, those living in a neighborhood with fewer than five enrolled households, those living in own home or subsidized housing, and those reporting work for landlord in lieu of rent.

DATA SOURCES: Initial Household Report Form, Housing Evaluation Form.

APPENDIX IV

DESIGN OF THE DEMAND EXPERIMENT

This appendix presents a brief overview of the Demand Experiment's purpose, reports, data collection, experimental design, and sample allocation.

IV.1 PURPOSE OF THE DEMAND EXPERIMENT

The Demand Experiment is one of three experiments established by the U.S. Department of Housing and Urban Development (HUD) as part of the Experimental Housing Allowance Program.[1] The purpose of these experiments is to test and refine the concept of housing allowances.

Under a housing allowance program, money (the allowance) is given directly to individual families in need to assist them in obtaining adequate housing. The allowance may be tied to housing by making the amount of the allowance depend on the amount of rent paid or by requiring that households meet certain housing requirements to receive the allowance payment. The initiative in using the allowance and the burden of meeting housing requirements are placed on the individual family rather than on developers, landlords, or the government.

The desirability, feasibility, and appropriate structure of a housing allowance program have not been established. Housing allowances could be less expensive than some other kinds of housing programs because they allow fuller utilization of existing sound housing; the allowance is not necessarily tied to new construction or to special classes of dwelling units. Housing allowances may also be more equitable. The allowance can be adjusted rapidly to changes in income without forcing the family to change units. Recipient families may, if they desire, use their own resources (by either paying higher rent or searching carefully) to obtain better housing than is required to receive the allowance. As long as program requirements are met, housing allowances permit families considerable choice in determining

[1] The other two experiments are the Housing Allowance Supply Experiment and the Administrative Agency Experiment.

the housing they want--where they live (near schools, near work, near friends, or relatives), or the type of unit they live in (single-family or multi-family). Finally, housing allowances could be less costly to administer. Program requirements need not cover every detail of participant housing. The burden of specifying and administering details that are not essential to the government, and of obtaining housing that meets requirements that are essential, is shifted from program administrators to participants and the private market. Because the program is less visible (the action in the housing market rests with individual families and can be dispersed over the entire market), there may be less public pressure on the administering agency.

These potential advantages are not unquestioned. Critics of housing allowances have suggested that poor families may lack the necessary experience with and knowledge of the private market for better housing to use allowances effectively; that special groups such as the elderly will not be effectively served without direct intervention to change the supply of housing to meet their needs; that administrative costs could rise uncontrollably; and that increasing the demand for housing without direct support for construction of new units will result in a substantial inflation of housing costs.

If housing allowances are desirable, they could be implemented by means of many different program structures. There is a wide range of possible allowance formulas, housing requirements, nonfinancial support (such as counseling), and administrative practices which could substantially affect both the costs and impact of a housing allowance program.

The Demand Experiment addresses issues of feasibility, desirability, and appropriate structure in terms of how individuals (as opposed to the market or administering agencies) react to various allowance formulas and housing standards requirements. The analyses and reports are designed to answer six policy questions:

1. <u>Participation</u>

   Who participates in a housing allowance program? How does the form of allowance affect the extent of participation for various households?

2. <u>Housing Improvements</u>

Do households receiving housing allowances in fact improve the quality of their housing? At what cost? How do households receiving a housing allowance seek to improve their housing--by moving, by rehabilitation? With what success?

3. <u>Locational Choice</u>

For those participants who move, how do the locational choices of allowance recipients compare with existing residential patterns? Are there nonfinancial barriers to effective use of a housing allowance?

4. <u>Administrative Issues</u>

What administrative issues and associated costs are involved in the implementation of a housing allowance program?

5. <u>Form of Allowance</u>

How do the different forms of a housing allowance compare in terms of participation, housing quality achieved, locational choice, costs (including administrative costs), and equity?

6. <u>Comparison with Other Programs</u>

How do housing allowances compare with existing housing programs and with income maintenance in terms of participation, housing quality achieved, locational choice, costs (including administrative costs), and equity?

The first three policy questions ask about the results of a housing allowance program. Participation can substantially affect both program costs and program desirability. Income transfer programs ordinarily do not enroll all those who are eligible. This obviously affects their potential scale and costs. At the same time, if a program fails to reach such key groups as the very poor, it may fail in its purpose, no matter how successful it is for those who do participate.

The issue of participation is particularly important in a housing allowance program. Such a program does not simply offer more money to needy households. It generally requires that they meet certain housing requirements to participate. The extent and nature of these requirements may make successful participation more or less difficult and desirable for various groups, such as the very poor, the elderly, or minorities.

The improvement in housing achieved under a housing allowance program is obviously central to judging its success. Housing improvement may be measured in terms of the change in the amount of housing purchased (essentially, the rent paid), achievement of certain specified quality levels in housing, or participant preferences and satisfaction with housing. Major issues include not only how these measures of housing change but what measures are most appropriate.

By providing poor households with a greater range of locational choice, a housing allowance may alter patterns of racial and socioeconomic segregation. In any case, the ability and interest of eligible households in searching for new housing can substantially affect their ultimate benefits from a housing allowance program. Examination of the degree of success with which households search for new housing may suggest the need for nonfinancial support, such as counseling, provision of vacancy lists, or equal opportunity support.

The fourth policy question concerns administrative issues. Although administrative issues are not a central concern of the Demand Experiment, analysis of the procedures used in the experiment may shed some light on selected issues, such as verification of participant income and household size, the need of providing housing information to participants, or appropriate coordination with other transfer programs.

The Demand Experiment studies a variety of potential housing allowance programs. It is designed to allow policymakers to make an informed choice among alternative forms of housing allowance programs. The fifth policy question asks how the effects of the allowance in terms of participation, housing change, locational choice, equity, and costs vary across different forms of housing allowance programs.

The last policy question asks how a housing allowance program compares with other housing programs or with income maintenance in terms of participation, housing quality achieved, locational choice, costs, and equity.

IV.2 REPORTS

The first analytic reports from the Demand Experiment will be submitted in 1976 and early 1977. These reports will examine key analytic issues using data collected during the first year of participation. They are intended to test basic analytic models and concepts and to identify areas for further work. The topics for these reports are grouped around areas defined by the first three policy questions: participation, housing consumption, and location.

The final set of reports, to be submitted in 1977 and 1978, will be based on the full two years of experimental data and will represent the final analytic products of the experiment. These reports address each of the six policy questions in turn.

IV.3 DATA COLLECTION

The Demand Experiment is conducted at two sites--Allegheny County, Pennsylvania (Pittsburgh), and Maricopa County, Arizona (Phoenix). Most of the information on participating households is collected from:

Baseline Interviews conducted by an independent survey operation before households are offered enrollment

Initial Household Report Forms and monthly Household Report Forms completed during and after enrollment to provide operating and analytic data on household size and income and on expenditures for housing

Supplements to the Household Report Forms completed after enrollment to provide data on assets, income from assets, actual taxes paid, income from self-employment, and extraordinary medical expenses

Housing Evaluation Forms completed by site office evaluators at least once each year for every dwelling unit occupied by participants, to provide information on the quality of participant housing

Periodic Interviews conducted approximately 6, 12, and 24 months after enrollment by an independent survey operation

Exit Interviews conducted by an independent survey operation for a sample of households that decline the enrollment offer or leave the program.

Surveys and housing evaluations are also administered to a sample of participants in existing housing programs.

The experimental programs in the Demand Experiment continue for three years after enrollment is completed. At the end of that time, eligible and interested allowance families will be aided in entering other housing programs, especially the Section 23 Leased Housing Program. Analysis will be based on data from only the first two years of participation. The experimental programs are continued for one additional year to avoid confusing participants' reactions to the ongoing experiment with their adjustments to the phaseout of the experiment.

IV.4 ALLOWANCE PLANS USED IN THE DEMAND EXPERIMENT

The Demand Experiment directly tests three combinations of payment formulas and housing requirements and five to six variations within each of these combinations--a total of 17 variations. These 17 variations allow some possible program designs to be tested directly. More important, they allow estimation of key responses in terms of such basic program parameters as the level of allowances, the level and type of housing requirements, the minimum fraction of its own income which the family is expected to contribute toward housing, and the way in which allowances vary with family size, income, and rent. These response estimates can then be used to address the policy questions, not just for the program plans directly tested but for a much larger set of candidate program plans.[1]

---

[1]The basic design and analysis approach, as approved by the HUD Office of the Policy Development and Research, is presented in Abt Associates Inc., Experimental Design and Analysis Plan of the Demand Experiment, Cambridge, Mass., March 1973, revised August 1973, and in Abt Associates Inc., (footnote continued on next page)

Two payment formulas are used in the Demand Experiment--Housing Gap and Percent of Rent.

Under the Housing Gap formula, payments to families constitute the difference between a basic payment level, C, and some reasonable fraction of family income. The payment formula is

$$P = C - bY$$

where P is the payment amount, C is the basic payment level, "b" is the rate at which the allowance is reduced as income increases, and Y is the net family income.[1] In the experiment, the basic payment level, C, varies with household size and is proportional to C*, the estimated cost of modest, existing standard housing at each site, and varies by household size.[2] Thus, the payment in the Housing Gap formula can be interpreted as making up the difference between some fraction of the cost of decent housing and the fraction of its own income that a household should be expected to pay for housing.

Under the Percent of Rent formula, the payment is a percentage of the family's rent. Thus, the payment is determined by

$$P = aR$$

where R is rent and "a" is the fraction of rent paid by the allowance. The values of "a" remain constant once a family has been enrolled.[3]

(footnote continued from previous page)
Summary Evaluation Design, Cambridge, Mass., June 1973. Details of the operating rules of the Demand Experiment are contained in Abt Associates Inc., Site Operating Procedures Handbook, Cambridge, Mass., April 1973, updated periodically.

[1] In addition, whatever the payment calculated by the formula, the actual payment cannot exceed the rent paid.

[2] For more detailed discussion regarding the derivation of C*, refer to Abt Associates Inc., Working Paper on Early Findings, Cambridge, Mass., January 1975, Appendix II.

[3] Five values of "a" are used in the Demand Experiment. Once a family is assigned its "a" value, the value generally stays constant in order to aid experimental analysis. In a national Percent of Rent program, "a" would probably vary with income and/or rent. Even in the experiment, if a family's income rises beyond a certain point, the "a" drops rapidly to zero. Similarly, the payment under Percent of Rent cannot exceed C* (the maximum payment under the modal Housing Gap plan); this effectively limits the rent subsidized to rents less than C*/a.

The Percent of Rent payment formula is directly tied to rent: a household's allowance payment is proportional to the total rent. Under the Housing Gap formula, however, two additional housing requirements are needed to tie the allowance to housing: Minimum Standards and Minimum Rent.

Under the Minimum Standards requirement, participants must occupy dwellings meeting certain standards to receive the allowance payment. Participants occupying units that do not meet these standards must either move or arrange to improve their current units to meet the standards. Participants already living in housing that meets standards may use the payment to pay for better housing or to reduce their rent burden (the fraction of income spent on rent) in their existing units.

If housing quality were broadly defined to include all residential services, and if rent levels were highly correlated with the level of services, then a straightforward housing requirement (one relatively inexpensive to administer) would be that recipients spend some minimum amount on rent. Minimum Rent is considered as an alternative to Minimum Standards in the Demand Experiment, so that differences in response and cost may be observed and the relative merits of the two types of requirements assessed. Although the design of the experiment uses a fixed minimum rent for each household size, a program for direct cash assistance could employ more flexible versions. Such versions could, for example, combine features of the Percent of Rent formula with the Minimum Rent requirement.[1] Thus, the three combinations of payment formulas and housing requirements used in the Demand Experiment are Housing Gap Minimum Standards, Housing Gap Minimum Rent, and Percent of Rent.

The Housing Gap allowance plans are shown in Table IV-1 below. The first nine plans all have "b" equal to 0.25, and include three variations in the level of C (1.2C*, C*, and 0.8C*) and three variations in housing requirements (Minimum Standards, Minimum Rent Low (0.7C*) and Minimum Rent High (0.9C*)). The next two plans have the same level of C (C*) and the Minimum Standards Housing Requirement, but different levels of "b"--the tenth plan

---

[1] For example, instead of receiving nothing if their rent is less than the Minimum Rent, households might be paid a fraction of their allowance depending on the fraction of Minimum Rent paid.

Table IV-1

HOUSING GAP ALLOWANCE PLANS

Housing Gap Formula: $P = C - bY$ where C is a multiple of C*

b VALUE	C LEVEL	HOUSING REQUIREMENTS			
		MINIMUM STANDARDS	MINIMUM RENT LOW = 0.7C*	MINIMUM RENT HIGH = 0.9C*	NO REQUIREMENT
b = .15	C*	Plan 10			
b = .25	1.2C*	Plan 1	Plan 4	Plan 7	
	C*	Plan 2	Plan 5	Plan 8	Plan 12
	0.8C*	Plan 3	Plan 6	Plan 9	
b = .35	C*	Plan 11			

Symbols:

b = Rate at which the allowance decreases as the income increases.

C* = Basic payment level (varied by family size and also by site).

has "b" equal to 0.15 while the eleventh plan has "b" equal to 0.35. The twelfth plan has no housing requirement.

Eligible households that do not meet the housing requirement can still enroll. They receive full payments whenever they meet the requirements and may do so anytime during the three years of the experiment. Even before they meet the housing requirements, such households receive a payment of $10 per month if they complete all reporting and interview requirements.

Within the Housing Gap design, the mean effects of changes in the allowance level and housing requirement can be estimated for all major responses. In addition, interactions between allowance level and housing requirement can be assessed. Responses to variations in the allowance/income schedule (changes in "b") can be estimated for the basic combination of the Minimum Standards housing requirement and C*.

The Percent of Rent allowance plans consist of five variations in "a", the proportion of rent paid to the household, as shown in Table IV-2 below.[1]

Table IV-2

PERCENT OF RENT ALLOWANCE PLANS

Percent of Rent Payment Formula: $P = aR$

Allowance Plan	13	14-16	17-19	20-22	23
Value of "a"	0.6	0.5	0.4	0.3	0.2

A demand function for housing will be estimated primarily from the Percent of Rent observations. This demand function should provide a powerful tool for analysis of alternative forms and parameter levels of housing allowance programs.

In addition to the various allowance plans, Control groups are necessary to establish a reference level for household responses, because a number of

[1] Designation of multiple plans for certain "a" values reflects an early assignment convention and does not indicate that the households in these plans are different.

uncontrolled factors may also induce changes in family behavior during the course of the experiment. Control households receive a monthly cooperation payment of \$10. They report the same information required of households receiving allowance payments, including household composition and income; they permit housing evaluations; and they complete the Baseline Interview and the three Periodic Interviews. (Control households are paid an additional \$25 fee for each Periodic Interview.)

Two Control groups are used in the Demand Experiment. Members of one group (Plan 24) were offered a Housing Information Program when they joined the experiment, and were paid \$10 for each of five sessions attended. (This program was also offered to all households that were offered allowances, but these households were not paid for attending sessions.) The other Control group (Plan 25) was not offered the Housing Information Program.

All the households in the various allowance plans had to meet a basic modal income eligibility requirement. This was defined (approximately) by the income level at which the household would receive a zero payment under the Housing Gap formula:

$$P = C^* - 0.25Y.$$

In addition, households in plans with lower payment levels (Plans 3, 6, 9, and 11) had to have incomes low enough to receive payments under these plans. Finally, only households with incomes in the lower third of the eligible population were eligible for enrollment in Plan 13 and only those in the upper two thirds were eligible for Plan 23.

IV.5 THE SAMPLE AFTER TWO YEARS

Much of the analysis of the impact of the housing allowance will be based on two years of experimental data. For this report the primary sample is all enrolled households (with exclusions specified in tables); Section 4.3 also uses the two-year sample. Table IV-3 presents the sample sizes for households active at enrollment and after two years for each treatment plan.

Active households include both households receiving a full payment and those not receiving a full payment. Households receiving full payments meet all

Table IV-3

SAMPLE SIZE AT ENROLLMENT AND TWO YEARS AFTER ENROLLMENT BY ALLOWANCE PLANS

ALLOWANCE PLAN[a]	ENROLLMENT SAMPLE		TWO YEAR SAMPLE	
	PITTSBURGH	PHOENIX	PITTSBURGH	PHOENIX
TOTAL HOUSING GAP	701	765	512	421
1	43[b]	48	33	30
2	59	74	42	35
3	62	66	43	39
4	43	42	34	24
5	62	70	50	39
6	61	63	44	35
7	45	43	30	30
8	67	78	44	44
9	67	70	43	35
10	57	64	45	36
11	60	77	41	34
12	75	70	63	40
TOTAL PERCENT OF RENT	510	490	407	298
13	34	32	28	21
14-16	121	114	109	81
17-19	145	120	113	66
20-22	118	140	92	84
23	92	84	65	46
TOTAL CONTROL[b]	434	525	321	282
24	210	262	159	137
25	224	263	162	145
TOTAL	1645	1780	1240	1001

SAMPLE: Enrollment Sample: All enrolled households, excluding those above the income eligibility limit.
Two Year Sample: Households active at two years, excluding those enrolled above the income eligibility limit, and those that moved into subsidized housing or their own homes.

DATA SOURCE: Payments file.

a. See Tables IV-1 and IV-2 for a description of the allowance plans.

b. Control households in plan 24 were offered the Housing Information Program; those in plan 25 were not.

requirements (including the housing requirements) and receive the full subsidy for which they are eligible given income, household size, and rent. Those not receiving a full payment receive only a monthly cooperation payment. Households fall in the latter group if they are homeowners, live in subsidized housing, have not met housing requirements, or have not turned in a rent receipt, but at the same time meet all other reporting and eligibility requirements.

APPENDIX V

PROGRAM HOUSING AND OCCUPANCY MEASURES

This appendix describes the housing and occupancy measures used in the analysis. These measures are based on the Minimum Standards housing requirement used in one part of the experiment. While such a requirement is not imposed on the Percent of Rent households included in the analysis in this report, they provide a convenient, if crude, measure of dwelling unit quality.

The discussion is organized as follows. The first section discusses the derivation of the program Minimum Standards requirement for housing quality, with special attention to its relationship to the American Public Health Association (APHA) code. The second section describes the various components of the Minimum Standards requirement including both physical requirements on the dwelling unit and occupancy requirements.

V.1 DEVELOPMENT OF THE MINIMUM STANDARDS HOUSING REQUIREMENT

There being no specific, generally accepted definition of standard housing, the program definition of minimum standards, which includes housing and occupancy standards, was not predetermined and thus had to be developed.

The American Public Health Association-Public Health Service (APHA-PHS) Recommended Housing Maintenance and Occupancy Ordinance (revised 1971) code and the Urban Institute's modification of it[1] served as the basic model for defining the standards. Table V-1 shows the relationship between this model and the program standards. The table compares the elements of the APHA Code, the Urban Institute's modification, and the Minimum Standards requirement. An element is indicated as comparable if the general meaning is similar, even though it may not be treated identically by all three.

---

[1] Urban Institute Working Paper No. 205-8, April 28, 1972.

Table V-1

POTENTIAL ELEMENTS FOR HOUSING STANDARDS

	APHA[1] CODE	APHA/CODE MODIFIED BY UI[2]	MINIMUM STANDARDS PROGRAM DEFINITION
Space per Occupant			
Total space	X	X	(3)
Max # persons per room or per bedroom	X	X	X
INTERIOR STRUCTURE			
Closet space	X	X	(1)
Exits	X	X	X[3]
Walls and Ceilings	X	X	X
Ceiling Height	X	X	X
Floors	X	X	X
Stairways	X	X	(1)
Ext. doors, skylights	X	X	(4)
Windows	X	X	Included under Ventilation
HEATING, ELECTRICITY, VENTILATION			
Electrical outlets	X	X	X
Heating	X	X	X
Venting (of heating)	X	X	X
Ventilation (windows)	X	X	X
OTHER STRUCTURAL REQUIREMENTS			
Handrails	X	X	(1)
Rat proofing	X	X	(3)
Screens on low windows	X		(3) (1)
Rat proofing, ext. doors, openings	X		(3)
Concrete basement floor	X		(3)
Rat proof basement walls	X		(3)
OUTSIDE CONDITIONS			
Trash and refuse	X	X	(1)

Key

Reasons for not including element in Minimum Standards Program Definition:

(1) Too stringent
(2) Too infrequent
(3) Too complicated or time consuming to evaluate
(4) Subsumed by other measure

1. American Public Health Association
2. Urban Institute
3. Removed as requirement effective November, 1973.

Table V-1 (continued)

	APHA CODE[1]	APHA/CODE MODIFIED BY UI[2]	MINIMUM STANDARDS PROGRAM DEFINITION
EXTERIOR			
Fences	X		(2)
Accessory structures	X		(2)
Foundation	X		(3)
Roof structure	X		X
Stairs/Porches	X		(1)
Plumbing & Installation	X		Plumbing facilities rated instead of installation
Chimneys and flues	X		(1)
Fire proof const. (local ordinance)	X		(3)
Wall structure			X
Wall surfaces			X
KITCHEN			
Stove	X	X	X
Refrigerator	X	X	X
Sink w/hot & cold water	X	X	X[3]
Counter & Cabinets	X	X	(1)
Complete kitchen facilities			X
Ceiling or wall-type light fixture	X		X
BATHROOM			
Flush toilet	X	X	X
Bathroom sink	X	X	X
Shower/tub	X	X	X
Ventilation	X	X	X
Bathroom door	X	X	(4)
Drug storage facility	X	X	(1)
Ceiling or wall-type light fixture	X	X	X

Key

Reasons for not including element in Minimum Standards Program Definition:

(1) Too stringent
(2) Too infrequent
(3) Too complicated or time consuming to evaluate
(4) Subsumed by other measure

1. American Public Health Association
2. Urban Institute
3. Removed as requirement effective November, 1973.

V.2 DESCRIPTION OF MINIMUM STANDARDS HOUSING REQUIREMENT

Table V-2 is a list of the Minimum Standards in the housing requirement as they apply to the dwelling unit itself. The requirements are grouped into 15 components made up of related items.

Occupancy requirements are separate from the physical requirements listed in Table V-2. However, the requirements for light-ventilation, ceiling height, and electrical service are applied to bedrooms in determining the number of adequate bedrooms for the program occupancy requirement as explained below.

The occupancy requirement sets a maximum of two persons for every adequate bedroom, regardless of age. (A studio or efficiency apartment is counted as a bedroom for occupancy standards.) An adequate bedroom is a room that can be completely closed off from other rooms and that meets the following program housing standards: ceiling height, light/ventilation, and electrical services. In addition, the room must meet the housing standards for the condition of room structure, room surface, floor structure, and floor surface.

Roomers and boarders are added to household size when determining whether a household meets occupancy standards, because all of the rooms in the dwelling unit are taken into account.

Table V-2

COMPONENTS OF MINIMUM STANDARDS
(Program Definition)

1. COMPLETE PLUMBING

Private toilet facilities, a shower or tub with hot and cold running water, and a washbasin with hot and cold running water will be present and in working condition.

2. COMPLETE KITCHEN FACILITIES

A cooking stove or range, refrigerator, and kitchen sink with hot and cold running water will be present and in working condition.

3. LIVING ROOM, BATHROOM, KITCHEN PRESENCE

A living room, bathroom, and kitchen will be present. (This represents the dwelling unit "core," which corresponds to an efficiency unit.)

4. LIGHT FIXTURES

A ceiling or wall-type fixture will be present and working in the bathroom and kitchen.

5. ELECTRICAL

At least one electric outlet will be present and operable in the living room and kitchen. A working wall switch, pull-chain light switch or additional electrical outlet will be present in the living room.[a]

6. HEATING EQUIPMENT

Units with no heating equipment; with unvented room heaters which burn gas, oil, or kerosene; or which are heated mainly with portable electric room heaters will be unacceptable.

---

a. This housing standard is applied to bedrooms in determining the number of adequate bedrooms for the program occupancy standard.

Table V-2 (continued)

7. ADEQUATE EXITS

There will be at least two exits from the dwelling unit leading to safe and open sapce at ground level. (For multi-family building only.) Effective November, 1973, (retroactive to program inception) this requirement was modified to permit override on case-by-case basis where it appears that fire safety is met despite lack of a second exit.

8. ROOM STRUCTURE

Ceiling structure or wall structure for all rooms must not be in condition requiring replacement (such as severe buckling or leaning).

9. ROOM SURFACE

Ceiling surface or wall surface for all rooms must not be in condition requiring replacement (such as surface material loose, containing large holes, or severely damaged).

10 CEILING HEIGHT

For living room, bathroom, and kitchen the ceiling must be 7 feet (or higher) in at least one-half of the room area.[a]

11. FLOOR STRUCTURE

Floor structure for all rooms must not be in condition requiring replacement (such as severe buckling or noticeable movement under walking stress).

12. FLOOR SURFACE

Floor surface for all rooms must not be in condition requiring replacement (such as large holes or missing parts).

---

a. This housing standard is applied to bedrooms in determining the number of adequate bedrooms for the program occupancy standard.

Table V-2 (continued)

13. ROOF STRUCTURE

The roof structure must be firm.

14. EXTERIOR WALLS

The exterior wall structure or exterior wall surface must not need replacement. (For structure this would include such conditions as severe leaning, buckling or sagging and for surface conditions such as excessive cracks or holes.)

15. LIGHT/VENTILATION

The unit will have a 10 percent ratio of window area to floor area and at least one openable window in the living room, bathroom and kitchen or the equivalent in the case of properly vented kitchens and/or bathrooms.[a]

---

a. This housing standard is applied to bedrooms in determining the number of adequate bedrooms for the program occupancy standard.

APPENDIX VI

MASTER VARIABLE LIST

This appendix contains a master list of variables tested in the process of building an hedonic index of housing. The first section describes the dependent variable, analytic rent. Sections 2 through 10 contain the independent variables describing tenure conditions and dwelling unit and neighborhood attributes. Each section lists the variables, their definitions, and their use in the final equations (i.e., included or excluded).

VI.1 ANALYTIC RENT (XACRA61H)

The basic definition of analytic rent used in the hedonic analysis is monthly payment for an unfurnished dwelling unit including basic utilities. The adjustments made in deriving analytic rent (XACRA61H) are summarized below:

Contract Rent

Contract rent is adjusted to a monthly amount to provide a common rental period.

Utilities Adjustment

Adjustments are made via site-specific tables for electricity, gas, heat, water, and trash collection if these are not included in contract rent. No adjustment is made for any other utilities or services, such as parking. The utility adjustment schedules were updated in February 1975. Both the original and revised schedules are contained in Table VI-1.

Furnishings Adjustment

For furtnished units a deduction for the cost of furnishings is made, equal to 11.5 percent of monthly contract rent adjusted for utilities. The analytical adjustment formula was developed from the 1974 Phoenix Housing Cost Panel's estimates of the additional cost of furnishings for units of varying sizes.[1] The

[1] The Pittsburgh panel's estimates were not used because they did not realistically reflect the additional cost of furnished units.

Table VI-1

UTILITY COST TABLES

(Dollar increment to contract rent per reported utility by size of dwelling unit)

NUMBER OF ROOMS IN DWELLING UNIT[c] (ANALYSIS DEFINITION)	ORIGINAL SCHEDULES[a]					REVISED SCHEDULES[b]				
	1,2	3	4	5	6+	1,2	3	4	5	6+
PITTSBURGH										
Electricity	$ 5	6	7	9	11	$ 5	9	10	12	14
Gas	2	2	3	3	4	2	3	3	4	5
Heating fuel	10	12	15	18	20	10	15	18	21	23
Garbage collection	3	3	3	3	3	3	3	3	3	3
Water	3	4	6	7	8	3	4	6	7	8
PHOENIX										
Electricity	$11	16	20	24	29	$13	18	23	28	33
Gas	5	6	7	11	15	6	7	8	12	17
Heating fuel	0	0	0	0	0	0	0	0	0	0
Garbage collection	3	3	3	3	3	3	3	3	3	3
Water	4	4	5	6	9	4	4	5	6	9

SOURCE: Local service and utility companies and public officials.

a. Effective through the month of January 1975

b. Effective from February 1975 forward.

c. Number of rooms is defined as number of rooms useable as living space (excluding bathrooms, half-rooms, unfinished basements or attics).

d. All refrigeration and air-conditioning costs are reflected in the table entries for electricity and gas.

rate of change in rent when nonfurnished and furnished units of the same unit size were compared was estimated to be an increase of 13 percent.[1] The estimate is transformed to an adjustment in the following manner:

let ACR = monthly contract rent + utilities adjustment
f = furnished units
n = nonfurnished units

$$ACR_n + .13ACR_n = ACR_f$$

or, $$1.13ACR_n = ACR_f$$

Solving for ACR_n: $$ACR_n = \frac{1}{1.13} ACR_f$$

$$ACR_n = .885ACR_f$$

The adjustment is not varied according to number of rooms or number of furnished rooms as it is assumed that the rent would reflect this.

The derivation of analytic rent, XACRA61H is therefore:

for unfurnished units: XACRA61H = Monthly Contract Rent + Utilities

for furnished units: XACRA61H = .885 (Monthly Contract Rent + Utilities)

or (Monthly Contract Rent + Utilities) - .115 (Monthly Contract Rent + Utilities)

Missing Rent Data

Analytic rent is missing for households that have missing contract rent data, that are "no cash renters," or that receive a reduction in rent because they work for the landlord but do not know the amount of the reduction. These households are excluded from the analysis.

[1]The mean of the incremental furnishings cost estimates was divided by the panel's estimated rental cost for each unit size and weighted by the actual distribution of unit sizes in the enrolled population. This produced an estimated percentage markup, for the rents of unfurnished units, of 13 percent.

VI.2 TENURE CONDITIONS

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XOCCRM	Number of persons per room. Number of persons divided by number of rooms, including kitchen and bathroom.	Pittsburgh, Phoenix, and common site equation
XRELATED	Related to the landlord. A 0-1 dummy variable (1 if any member of household is related to landlord).	Pittsburgh and Phoenix
XLLBLG	Landlord lives in the building. A 0-1 dummy variable (1 if landlord lives in the building).	Pittsburgh
XCONTACT	Number of landlord contacts for maintenance (e.g., rats, roaches, lack of heat, noise made by other tenants' children).	Pittsburgh, Phoenix, and common site equation
XLLVSAT	High satisfaction with landlord's response to requests for repairs. A 0-1 dummy variable (1 if tenant is very satisfied with landlord's response).	Not included in final equations
XLLSSAT	Satisfaction with landlord's response to requests for repairs. A 0-1 dummy variable (1 if tenant is somewhat satisfied with landlord's response).	Not included in final equations
XLLSDIS	Dissatisfaction with landlord's response to requests for repairs. A 0-1 dummy variable (1 if tenant is somewhat dissatisfied with landlord's response).	Not included in final equations
XLLVDIS	High dissatisfaction with landlord's response to requests for repairs. A 0-1 dummy variable (1 if tenant is very dissatisfied with landlord's response).	Not included in final equations
XLLWORK2	Landlord's maintenance of dwelling unit. Tenant's rating of landlord's maintenance (e.g., cleaning of grounds, extermination of rats and roaches, repair of appliances, general repair and painting of exterior of unit). Continuous scale--0 if landlord does not do anything, 16 if landlord maintains everything well.	Not included in final equations

VI.2 TENURE CONDITIONS continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XRES6MO	Residence 6 months or less. A 0-1 dummy variable (1 if tenant has occupied dwelling unit for six months or less).	Not included in final equations
XRES15	Residence 1 to 5 years. A 0-1 dummy variable (1 if tenant has occupied dwelling unit for 1 to 5 years).	Not included in final equations
XRES510	Residence 5 to 10 years. A 0-1 dummy variable (1 if tenant has occupied dwelling unit for 5 to 10 years).	Not included in final equations
XRESG10	Residence more than 10 years. A 0-1 dummy variable (1 if tenant has occupied dwelling unit for more than 10 years).	Not included in final equations
XLINGER	Length of residence. A continuous variable indicating length of time tenant has occupied the dwelling unit (months).	Not included in final equations
XLNLING	Natural log (ln) of length of residence. The natural logarithm of the length of residence.	Included in final common site equations
XEXP4	Length of residence. An exponential function indicating the length of residence $\left(XEXP4 = 1 - \exp\left(\frac{-XLINGER+12}{36}\right)\right)$	Pittsburgh
XEXP366	Length of residence. An exponential function indicating the length of residence $\left(XEXP366 = 1 - \exp\left(\frac{-XLINGER+6}{36}\right)\right)$	Phoenix

VI.3 DWELLING UNIT FEATURES

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XAPPL	Dishwasher and/or disposal provided. A 0-1 dummy variable (1 if dishwasher and/or disposal are provided).	Pittsburgh, Phoenix, and common site equation
XSTOREF	Stove or refrigerator provided. A 0-1 dummy variable (1 if stove or refrigerator are provided by the landlord at no extra cost to tenant).	Phoenix
XSTAREF	Stove and refrigerator provided. A 0-1 dummy variable (1 if stove and refrigerator are provided by the landlord at no extra cost to the tenant).	Pittsburgh and common site equation
XLLREFP	Refrigerator provided. A 0-1 dummy variable (1 if refrigerator is provided by the landlord at no extra cost to the tenant).	Not included in final equations
XLLSTOP	Stove provided. A 0-1 dummy variable (1 if a stove is provided by the landlord at no extra cost to the tenant).	Not included in final equations
XWINCOV	Window coverings. A 0-1 dummy variable (1 if most windows have screens (Phoenix) and 1 if most windows have screens in summer and storm windows in winter (Pittsburgh)).	Not included in final equations
XACPITT	Air-conditioning present, Pittsburgh. A 0-1 dummy variable (1 if unit has central air-conditioning or individual air-conditioning units).	Pittsburgh
XCACPHX	Central air-conditioning present, Phoenix. A 0-1 dummy variable (1 if unit has central air-conditioning).	Phoenix
XCOOLPHX	Other air-conditioning in Phoenix. A 0-1 dummy variable (1 if unit has an evaporation cooler or individual air-conditioning units).	Not included in final equations
XPARK	Parking facilities provided. A 0-1 dummy variable (1 if dwelling unit is provided with parking facilities).	Included in final common site equation

VI.3 DWELLING UNIT FEATURES continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XCARPRT	Carport provided. A 0-1 dummy variable (1 if dwelling unit is provided with a carport at no extra cost to tenant).	Not included in final equations
XGAR	Garage provided. A 0-1 dummy variable (1 if dwelling unit is provided with a garage at no extra cost to tenant).	Pittsburgh
XCARGAR	Carport or garage provided. A 0-1 dummy variable (1 if dwelling unit is provided with a carport or garage at no extra cost to tenant).	Phoenix
XOFFSTR	Offstreet parking provided. A 0-1 dummy variable (1 if dwelling unit is provided with off-street parking at no extra cost to tenant).	Pittsburgh
XPAINT	Recent interior painting or wallpapering. A 0-1 dummy variable (1 if any walls or ceilings have been painted or wall-papered within the past year and if some of the materials and labor were provided by the landlord).	Pittsburgh, Phoenix, and common site equation
XECLOSET	Enough closets. A 0-1 dummy variable (1 if tenant is satisfied with the number of closets provided with the unit).	Not included in final equations
XESTORAG	Enough storage space. A 0-1 dummy variable (1 if tenant is satisfied with the amount of storage space provided with the unit).	Not included in final equations
XYARD	Private yard. A 0-1 dummy variable (1 if unit has a yard and the yard is not shared).	Pittsburgh and common site equation
XNOYARD	No yard. A 0-1 dummy variable (1 if no yard is present).	Not included in final equations
XCENH	Central heat present. A 0-1 dummy variable (1 if dwelling unit has central warm air heating).	Phoenix

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XOTHH	Other types of heat. A 0-1 dummy variable (1 if dwelling unit has steam heat, or built-in electric units, or floor, wall, or pipeless furnace).	Not included in final equations
XBADH	Inferior or no heat. A 0-1 dummy variable (1 if dwelling unit has only room heaters without flue, or fireplace or stove, or portable electric heaters, or no heating equipment).	Pittsburgh, Phoenix, and common site equation
XHEATW	Working condition of heating system. A 5-point rating of heating system (0 if no rating is available, 4 if system is in good working condition).	Not included in final equations
XTEMP	Temperature control. A 0-1 dummy variable (1 if Pittsburgh unit is provided with central heat or air-conditioning; 1 if Phoenix unit is provided with central heat or central air-conditioning).	Included in final common site equation
XEHEAT	Adequate heat. A 0-1 dummy variable (1 if tenant is satisfied with the heat in winter).	Not included in final equations
XEHOTWTR	Enough hot water. A 0-1 dummy variable (1 if tenant is satisfied with the amount of hot water provided with the unit).	Not included in final equations
XKITCHP	Presence of adequate kitchen facilities. A 0-1 dummy variable (1 if dwelling unit has complete kitchen facilities as defined for Minimum Standard Housing). See Appendix V.	Pittsburgh
XKITCHOK	High quality kitchen. A 0-1 dummy variable (1 if unit has many high quality kitchen features: all kitchen facilities present and working; sink condition is "like new"; at least 4 sq. feet of counter space and 10 sq. feet of shelving exists; and all ratings for surface and structural quality of walls, ceilings and floor equal 0 ("like new") or 1 ("need only minor repairs")).	Pittsburgh
XPLUMP	Plumbing present. A 0-1 dummy variable (1 if unit is provided with private toilet facilities, a shower/tub and washbasin with hot/cold running water).	Phoenix

VI.3 DWELLING UNIT FEATURES continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XPLUMW	Working condition of plumbing. A 5-point rating of toilets (0 if no rating is available, 4 if toilets are in good working condition).	Pittsburgh
XHPLUMR	Presence of adequate plumbing. A 0-1 dummy variable (1 if unit has adequate plumbing facilities as defined for Minimum Standards Housing). See Appendix V.	Included in final common site equation
XBASOK	Basic Plumbing Facilities. A 0-1 dummy variable (1 if unit has both adequate plumbing and adequate kitchen facilities as defined for Minimum Standard Housing (see Appendix V) and no additional plumbing problem was cited by the evaluator on the Housing Evaluation Form).	Not included in final equations
XBATHOK	High quality bathroom. A 0-1 dummy variable (1 if unit has high quality bathroom features: shower/tub condition equal 1 ("like new") and both walls and floors are waterproof).	Not included in final equations
XHELECR	Presence of adequate electrical outlets. A 0-1 dummy variable (1 if unit has adequate electrical facilities as defined for Minimum Standard Housing). See Appendix V.	Not included in final equations
XEOUTLTS	Enough electrical outlets. A 0-1 dummy variable (1 if tenant is satisfied with the number of electrical outlets provided with the unit).	Not included in final equations
XHLIVER	Presence of adequate light and ventilation. A 0-1 dummy variable (1 if dwelling unit has adequate light and ventilation as defined for Minimum Standard Housing). See Appendix V.	Phoenix and common site equation
XEVENTIL	Enough ventilation. A 0-1 dummy variable (1 if tenant is satisfied with the ventilation provided in the unit).	Not included in final equations

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XHCEHTR	Presence of adequate ceiling height. A 0-1 dummy variable (1 if dwelling unit has adequate ceiling height as defined for Minimum Standard Housing). See Appendix V.	Pittsburgh, Phoenix, and common site equation
XHADQEXR	Presence of adequate exits. A 0-1 dummy variable (1 if unit has adequate exits as defined for Minimum Standard Housing). See Appendix V.	Pittsburgh
XLEAK	Walls, roofs or ceilings leak. A 0-1 dummy variable (1 if any walls, roofs, or ceilings leak when it rains).	Not included in final equations
XFANCY1	High quality features. A 0-1 dummy variable indicating presence of many good quality dwelling unit facilities (1 if complete plumbing facilities and complete kitchen facilities are present, core rooms are present, central heat is present in Pittsburgh and central, steam, electric or pipeless furnace present in Phoenix, central air-conditioning present in Phoenix, kitchen sink rating is like new or good, and bathroom rating is high quality).	Not included in final equations
XFANCY2	Many high quality features. A 0-1 dummy variable indicating presence of many good quality dwelling unit facilities plus adequate surface and structural quality (equals 1 if XFANCY1=1 and all surface and structure ratings for walls, ceilings, and floors are 0 ("like new") or 1 ("need minor repairs only")).	Pittsburgh and common site equation

VI.4 DWELLING UNIT SIZE

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XTOTRMS	Total rooms. The total number of rooms in the dwelling unit, including kitchen and bathroom.	Not included in final equations
XLTOTRMS	Natural log (ln) total rooms. The natural logarithm of the total number of rooms in the dwelling unit.	Pittsburgh, Phoenix, and common site equation
XAREAPR	Area per room. The square feet per room (total number of square feet divided by number of rooms).	Not included in final equations
XLAREAPR	Natural log (ln) area per room. The natural logarithm of the square feet per room.	Pittsburgh, Phoenix, and common site equation

VI.5 INTERIOR DWELLING UNIT QUALITY

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XRATINGR	Overall evaluator rating. The overall evaluator rating of the dwelling unit. Values range from 3 (best) to zero (worst).	Pittsburgh
XFAILS	Components failed. The number of components of Minimum Standard Housing which are not met. Refer to Appendix V.	Not included in final equations
XACSTR	Average ceiling structure. The average value of the four point ratings of ceiling structure in all rooms. Values range from 3 (best) to zero (worst).*	Included in overall average or in factor score
XACSUR	Average ceiling surface. The average value of the four point rating of ceiling surface in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAWSTR	Average wall structure. The average value of the four point rating of wall structure in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAWSUR	Average wall surface. The average value of the four point rating of wall surface in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAFSTR	Average floor structure. The average value of the four point rating of floor structure in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score

*The ratings for interior surface and structural quality, exterior quality, and window condition, used in dwelling unit quality variables, are based on four point scales ranging from 3 to zero, where 3 means "like new," 2 indicates "needs some repair," 1 indicates "needs substantial repair," and 0 indicates "needs replacement."

VI.5 INTERIOR DWELLING UNIT QUALITY continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XAFSUR	Average floor surface. The average value of the four point rating of floor surface in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XACEIL	Average ceiling quality. The average value of the ratings of surface and structural condition of the ceilings in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAWALL	Average wall quality. The average value of the ratings of surface and structural condition of the walls in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAFLOOR	Average floor quality. The average value of the ratings of surface and structural condition of the floors in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAWNCON	Window condition. The average value of the ratings of window condition (sash and panes) in all rooms. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XAEXT	Average exterior. The average of the four point ratings for exterior wall structure and exterior wall surface. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score
XTOTEXT	Total exterior. The average value of the ratings for exterior wall structure, exterior wall surface, roof surface, gutter condition, exterior stairs, exterior cleanliness, and roof structure. Values range from 3 (best) to zero (worst).	Included in overall average or in factor score

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XQUAL1	Average quality. The average of all ratings (for all rooms) of the surface and structural quality of ceilings, walls and floors. (XACEIL + XAWALL + XAFLOOR/3). Values range from 3 (best) to zero (worst).	Not included in final equations
XQUAL2	Average surface and structural quality. The average of all ratings on the surface and structural quality of floors, walls, and ceilings, of window condition, and of exterior quality. (XACEIL + XAWALL + XAFLOOR + XAWNCON + XTOTEXT)/5). Values range from 3 (best) to zero (worst).	Phoenix and common site equation

VARIABLES DERIVED FROM PRINCIPAL COMPONENTS ANALYSIS OF QUALITY RATINGS IN PITTSBURGH (SEE APPENDIX VII)

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XF1STR	Pittsburgh. Wall and ceiling structure. Factor score. Structural condition of walls and ceilings (living room, kitchen, bedroom, bathroom).	Not included in final equations
XF2SUR	Pittsburgh. Poor wall and ceiling surface. Factor score. Surface condition of walls and ceilings (living room, kitchen, bedroom).	Pittsburgh
XF3FSUR	Pittsburgh. Poor floor surface. Factor score. Surface quality of floors.	Not included in final equations
XF4WIN	Pittsburgh. Poor window condition. Factor score. Window condition.	Pittsburgh
XF5FSTR	Pittsburgh. Poor floor structure. Factor score. Structural quality of floors.	Not included in final equations
XF6BSUR	Pittsburgh. Poor bathroom wall and ceiling surface. Factor score. Surface condition of walls and ceilings in bathroom.	Pittsburgh

VI.6 BUILDING TYPE AND LOCATION

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XAGE	Building age. The age of the dwelling unit. (Age was coded in seven intervals: 1970 to present, 1960-69, 1950-59, 1940-49, 1930-39, 1920-29, 1919 or earlier. The values in years assigned to each of these intervals are, respectively, 3, 8, 18, 28, 38, 48, and 63 years).	Pittsburgh, Phoenix, and common site equation
XMULTI5	Large multifamily. A 0-1 dummy variable indicating large multifamily structures (1 if structure contains 5 or more units).	Pittsburgh, Phoenix, and common site equation
XSNGDET	Single-family detached. A 0-1 dummy variable (1 if structure is a single-family detached house).	Not included in final equations
XROWDUP	Row and duplex units. A 0-1 dummy variable (1 if structure is a row-house or duplex).	Not included in final equations
XMULT34	Small multifamily. A 0-1 dummy variable indicating smaller multifamily structures (1 if structure has 3 or 4 units).	Not included in final equations
XMOBILE	Mobile units. A 0-1 dummy variable (1 if unit is a mobile home).	Not included in final equations
XDIST	Distance from central business district A continuous variable indicating the distance from the central business district to the dwelling unit (in miles).	Phoenix and common site equation

VI.7 BLOCK FACE QUALITY

(The block face is defined as the area within 100 yards of the dwelling unit.)

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XLNDSCPR	Quality of block face landscaping. The quality of landscaping on the block face, rated on a scale of 0 to 3 (where 3 equals a full range of landscaping present and 0 represents no evidence of landscaping).	Phoenix and common site equation
XFANCYN	High quality block face. A 0-1 dummy variable, taking the value of 1 if many good quality block face features are present (good or excellent street maintenance, good or excellent landscaping, little or no litter, no abandoned buildings and no detrimental features, including noise, odor, physical hazards or flooding).	Pittsburgh
XSTRETMR	Street maintenance. Quality of street maintenance rated on a scale of 0 to 3 (where 3 represents well paved and maintained streets and 0 represents very poor maintenance).	Not included in final equations
XLITTERR	Street litter. Condition of street and sidewalks rated on a scale of 0 to 3 (where 3 represents clean streets and 0 represents a large accumulation of trash and litter on the block face).	Not included in final equations
XABAND	Abandoned buildings or cars. A 0-1 dummy variable (1 if abandoned buildings or abandoned cars are present on the block face).	Not included in final equations
XNNONE	No other residential structures. A 0-1 dummy variable (1 if no other residential structures are present).	Not included in final equations
XSTRQ	Average block face quality. The average of the 4 point ratings for quality of landscaping, quality of street maintenance, and presence of litter (XLNDSCPR, XSTRETMR, XLITTERR). A value of 3 indicates a high quality block face and a value of 0 indicates very poor quality.	Not included in final equations

VI.7 BLOCK FACE QUALITY continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XNUBEN	Beneficial features. The number of beneficial features in the immediate neighborhood (0 to 5) as indicated by evaluation entries for parklands, water, woodlands, view, and other unique features.	Not included in final equations

VI.8 CENSUS TRACT DESCRIPTORS

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XCTMDVAL	The median value of housing stock in the tract.	Included in factor score in Pittsburgh and Phoenix
XCTMDGR	The median value of gross rent in the tract.	Included in factor score in Pittsburgh and Phoenix
XCTOWNOC	The percent of owner-occupied units in tract.	Included in factor score in Pittsburgh and Phoenix
XCTDUAGE	The median age of housing stock in tract.	Included in factor score in Pittsburgh and Phoenix
XCTDUAC3	The percent of dwelling units with central air-conditioning.	Included in factor score in Pittsburgh and Phoenix
XCTSGFAM	Percent of single family detached units in tract.	Included in factor score in Pittsburgh and Phoenix
XCTMDSCH	Mean number of years of school of persons 25 and over.	Included in factor score in Pittsburgh and Phoenix
XCTMDINC	Median family income in tract.	Included in factor score in Pittsburgh and Phoenix, and included in final common site equation
XCTSMHSE	Percent in same houses since 1965 in tract.	Included in factor score in Pittsburgh and Phoenix
XCTBLCOL	Percent of blue collar (14 or more years) workers.	Included in factor score in Pittsburgh and Phoenix
XCTCRWD	Percent of dwelling units with 1.01 or more persons per room.	Included in factor score in Pittsburgh and Phoenix

VI.8 CENSUS TRACT DESCRIPTORS continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XCTGT4DU	Percent of dwelling units in tract with 5 or more units.	Included in factor score in Pittsburgh and Phoenix
XCTBATH	Percent of dwelling units with more than one bathroom.	Included in factor score in Pittsburgh and Phoenix
XCTMNRMS	Average number of rooms per dwelling unit in tract.	Included in factor score in Pittsburgh and Phoenix
XCTPBLK	Percent black in tract.	Included in factor score in Pittsburgh and Phoenix
XCTSPAM	Percent Spanish American in tract.	Included in factor score in Phoenix
XCTMDHSZ	Median household size in tract.	Included in factor score in Pittsburgh and Phoenix
XCTPTMGR	Percent of persons (14 or more years) who work in professional jobs.	Included in factor score in Pittsburgh and Phoenix
XCTCLSL	Percent of persons (14 or more years) who work in clerical/sales jobs.	Included in factor score in Pittsburgh and Phoenix
XCTWOPLB	Percent of dwelling units lacking 1 or more plumbing facilities.	Included in factor score in Pittsburgh and Phoenix
XCTWOHT	Percent of dwelling units lacking adequate heat.	Included in factor score in Pittsburgh and Phoenix
XCTSHKIT	Percent of dwelling units with incomplete or shared kitchen facilities.	Included in factor score in Pittsburgh and Phoenix
XCTCCSBR	Central city residence. A 0-1 dummy variable (1 if dwelling unit is in the city limits).	Not included in final equations

VI.8 CENSUS TRACT DESCRIPTORS continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XNACESS	Generalized employment accessibility. A generalized accessibility variable defined by a standard exponential decay function of travel time weighted by employment in each of 132 school districts in the Pittsburgh SMSA.	Not included in final equations

$$XNACESS = \sum_{j=1}^{n} \frac{E_j}{t_{ij}^{\alpha}}$$

where E_j = employment in zone j

t_{ij} = travel time from zone i to zone j

α = a positive constant.

Each school district is composed of several Census tracts; the data were allocated to Census tracts assuming no within-district variation in accessibility. The information was obtained from the National Bureau of Economic Research, Cambridge, Massachusetts.

VARIABLES DERIVED FROM PRINCIPAL COMPONENTS ANALYSIS OF CENSUS VARIABLES IN PITTSBURGH (SEE APPENDIX VII)

XCENF01	Pittsburgh. Census tracts with owner occupied, single family dwelling units. Factor score.	Not included in final equations
XCENF02	Pittsburgh. Census tracts with higher priced units and higher socioeconomic status.	Pittsburgh
XCENF03	Pittsburgh. Nonminority Census tracts with higher socioeconomic status. Factor score.	Pittsburgh

VI.8 CENSUS TRACT DESCRIPTORS continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XCENF04	Pittsburgh. Census tracts with blue collar workers and nonminority residents. Factor score.	Pittsburgh
XCENF05	Pittsburgh. Proportion Spanish American households in Census tract. Factor score.	Not included in final equations
XCTF01	Pittsburgh. Census tracts with owner-occupied single family units. Factor score. Excludes variables indicating racial composition of tracts.	Not included in final equations
XCTF02	Pittsburgh. Census tracts with higher socioeconomic status. Factor score. Excludes variables indicating racial composition of tracts.	Not included in final equations
XCTF03	Pittsburgh. Census tracts with newer, higher priced units. Factor score. Excludes variables indicating racial composition of tracts.	Not included in final equations
	VARIABLES DERIVED FROM PRINCIPAL COMPONENTS ANALYSIS OF CENSUS VARIABLES IN PHOENIX (SEE APPENDIX VII)	
XCENF01	Phoenix. Census tracts with higher priced units and higher socioeconomic status. Factor score.	Phoenix
XCENF02	Phoenix. Census tracts with owner-occupied single-family housing. Factor score.	Phoenix
XCENF03	Phoenix. Poor quality dwelling units in Census tract. Factor score.	Phoenix
XCTF01	Phoenix. Census tracts with higher priced units and higher socioeconomic status. Factor score. Excludes variables indicating racial composition of tracts.	Not included in final equations
XCTF02	Phoenix. Census tracts with owner-occupied single-family housing. Factor score. Excludes variables indicating racial composition of tracts.	Not included in final equations
XCTF03	Phoenix. Poor quality dwelling units in Census tract. Factor score. Excludes variables indicating racial composition of tracts.	Not included in final equations

VI.9 NEIGHBORHOOD VARIABLES

(Aggregation of Census tracts into larger neighborhoods, the C* neighborhoods)

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XHCNTREE	Quality of landscaping. The mean value of respondent ratings (4 point scale) on the quality of landscaping, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBTRE	Poor quality landscaping. The proportion of respondents in a C* neighborhood who rated the quality of landscaping as poor.	Not included in final equations
XHCNPARK	Quality of parking. The mean value of respondent ratings (4 point scale) on the availability of parking, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBPRK	Poor availability of parking. The proportion of respondents in a C* neighborhood that rated availability of parking as poor.	Not included in final equations
XHCNGARB	Quality of garbage collection. The mean value of respondent ratings (4 point scale) on the quality of garbage collection service, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNFIRE	Quality of fire department service. The mean value of respondent ratings (4 point scale) on the quality of fire protection, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNTRAN	Quality of public transportation. The mean value of respondent ratings (4 point scale) on the availability of public transportation facilities, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNPOL	Quality of police protection. The mean value of respondent ratings (4 point scale) on the quality of police protection, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBPOL	Poor quality police protection. The proportion of respondents in a C* neighborhood who rated the quality of police protection as poor.	Not included in final equation

VI.9 NEIGHBORHOOD VARIABLES continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XHCNMED	Quality of medical facilities. The mean value of respondent ratings (4 point scale) on the accessibility to medical facilities, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNLGHT	Quality of street lighting. The mean value of respondent ratings (4 point scale) on the quality of street lighting aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNCARE	Quality of daycare facilities. The mean value of respondent ratings (4 point scale) on the quality of daycare facilities, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNELM	Quality of elementary schools. The mean value of respondent ratings (4 point scale) on the quality of elementary schools, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBELM	Poor quality of elementary schools. The proportion of respondents in a C* neighborhood who rated the quality of elementary schools as poor.	Not included in final equations
XHCNJH	Quality of junior high schools. The mean value of respondent ratings (4 point scale) on the quality of junior high schools, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBJH	Poor quality junior high schools. The proportion of respondents in a C* neighborhood that rated junior high schools as poor.	Not included in final equations
XHCNHSCH	Quality of senior high schools. The mean value of respondent ratings (4 point scale) on the quality of senior high schools, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBSH	Poor quality senior high schools. The proportion of respondents in a C* neighborhood that rated senior high schools as poor.	Not included in final equations
XHCNPLAY	Quality of children's recreation. The mean value of respondent ratings (4 point scale) on the quality of play areas for children, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XHCNBRCK	Poor quality children's recreation. The proportion of respondents in a C* neighborhood that rated children's recreation as poor.	Not included in final equations
XHCNTREC	Quality of teenage recreation. The mean value of respondent ratings (4 point scale) on the quality of teenage recreation, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNBRCT	Poor quality teenage recreation. The proportion of respondents in a C* neighborhood that rated teenage recreation as poor.	Not included in final equations
XHCNAREC	Quality of adult recreation. The mean value of respondent ratings (4 point scale) on the availability of adult recreation facilities, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix and included in final common site equation
XHCNBRCA	Poor quality adult recreation. The proportion of respondents in a C* neighborhood who rated the quality of adult recreation facilities as poor.	Not included in final equations
XHCNPRAY	Convenience to places of worship. The mean value of respondent ratings (4 point scale) on the convenience to places of worship, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNSHOP	Convenience to grocery shopping. The mean value of respondent ratings (4 point scale) on the convenience to grocery shopping, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNCRIM	Crime problems. The mean value of respondent ratings (3 point scale) on the extent to which crime is a problem, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix

VI.9 NEIGHBORHOOD VARIABLES continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XHCNDRUG	Drug problems. The mean value of respondent ratings (3 point scale) on the extent to which drug users are a problem, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNLOUD	Problems with noise. The mean value of respondent ratings (3 point scale) on problems with noise, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNABAN	Abandoned buildings. The mean value of respondent ratings (3 point scale) on presence of abandoned buildings, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNSTR	Problems with street repair. The mean value of respondent ratings (3 point scale) on problems with street repair, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNMESS	Litter and trash problems. The mean value of respondent ratings (3 point scale) on problems with litter and trash, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNJUNK	Junk-filled lots. The mean value of respondent ratings (3 point scale) on the extent to which junk-filled lots are a problem, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNTRAF	Heavy traffic problems. The mean value of respondent ratings (3 point scale) on problems of heavy traffic, aggregated by C* neighborhood.	Included in factor score in Pittsburgh and Phoenix
XHCNQNBN	Neighborhood quality. The mean value of respondent ratings (4 point scale) on the quality of police protection, landscaping, elementary schools, and recreation facilities, aggregated by C* neighborhood.	Not included in final equations

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XHCNPNBD	Neighborhood problems. The mean value of respondent ratings (3 point scale) on the extent to which crime, vacant lots, litter, and trash are a problem aggregated by C* neighborhood.	Not included in final equations
XHCNBNBD	Poor quality neighborhood services. The average proportion of respondents who rated the quality of police services, recreation, schools, and landscaping as poor, aggregated by C* neighborhood.	Not included in final equations
	VARIABLES DERIVED FROM PRINCIPAL COMPONENTS ANALYSIS OF PARTICIPANT NEIGHBORHOOD RATINGS (AGGREGATED BY C* NEIGHBORHOOD) IN PITTSBURGH (SEE APPENDIX VII)	
XCNHF11	Pittsburgh. Good recreational facilities and access to shopping and transportation. Factor score.	Pittsburgh
XCNHF12	Pittsburgh. School quality. Factor score.	Not included in final equations
XCNHF13	Pittsburgh. Traffic and litter problems. Factor score.	Pittsburgh
XCNHF14	Pittsburgh. Problems with crime and public services. Factor score.	Pittsburgh
XCNHF21	Pittsburgh. Overall neighborhood problems. Factor score.	Not included in final equations
XCNHF22	Pittsburgh. Poor schools and police protection. Factor score.	Not included in final equations
XCNHF23	Pittsburgh. Absence of recreation facilities. Factor score.	Not included in final equations

VI.9 NEIGHBORHOOD VARIABLES continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
	VARIABLES DERIVED FROM PRINCIPAL COMPONENTS ANALYSIS OF PARTICIPANT NEIGHBORHOOD RATINGS (AGGREGATED BY C* NEIGHBORHOOD, IN PHOENIX (SEE APPENDIX VII)	
XCNHF11	Phoenix. Overall neighborhood quality. Factor score.	Phoenix
XCNHF12	Phoenix. Recreational facilities. Factor score.	Phoenix
XCNHF13	Phoenix. School quality and lack of transportation facilities. Factor score.	Not included in final equations
XCNHF14	Phoenix. Access to shopping and parking. Factor score.	Phoenix
XCNHF15	Phoenix. Traffic problems. Factor score.	Not included in final equations
XCNHF16	Phoenix. Fire protection and garbage service. Factor score.	Not included in final equations
XCNHF21	Phoenix. Absence of recreation facilities. Factor score.	Not included in final equations
XCNHF22	Phoenix. Problems with litter, abandoned cars, and police protection. Factor score.	Not included in final equations
XCNHF23	Phoenix. Poor elementary schools. Factor score.	Not included in final equations
XCNHF24	Phoenix. Drug and crime problems. Factor score.	Not included in final equations
XCNHF25	Phoenix. Traffic problems. Factor score.	Not included in final equations

VI.10 RACE OR ETHNICITY OF HOUSEHOLD AND RACIAL SUBMARKETS

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XBLACK	Black household. A 0-1 dummy variable (1 if head of household is black).	Variables used to test for racial price discrimination
XSPAN	Spanish American. A 0-1 dummy variable (1 if head of household is Spanish American).	Variables used to test for racial price discrimination
XGHETTO	Minority submarket. A 0-1 dummy variable (1 if the Census tract is greater than 50% black in Pittsburgh or 50% Spanish American in Phoenix).	Variables used to test for racial price discrimination
XMIXED	Racially mixed submarket. A 0-1 dummy variable (1 if Census tract is between 20% and 50% black in Pittsburgh or Spanish American in Phoenix).	Variables used to test for racial price discrimination
XWHITE	White submarket. A 0-1 dummy variable (1 if in the Census tract the proportion of black households in Pittsburgh and Spanish American households in Phoenix is less than 20%).	Variables used to test for racial price discrimination
XMGHET	Minority household in minority market. A 0-1 dummy variable (1 if minority household lives in minority tract--black in Pittsburgh, Spanish American in Phoenix).	Variables used to test for racial price discrimination
XWGHET	White household in minority market. A 0-1 dummy variable (1 if white household lives in minority tract).	Variables used to test for racial price discrimination
XMMIXED	Minority household in mixed market. A 0-1 dummy variable (1 if minority household lives in racially mixed tract).	Variables used to test for racial price discrimination
XWMIXED	White household in mixed market. A 0-1 dummy variable (1 if white household lives in racially mixed tract).	Variables used to test for racial price discrimination
XMWHITE	Minority household in white market. A 0-1 dummy variable (1 if minority household lives in white tract).	Variables used to test for racial price discrimination

VI.10 RACE OR ETHNICITY OF HOUSEHOLD AND RACIAL SUBMARKETS continued

ACRONYM	TITLE	USE IN FINAL HEDONIC EQUATIONS
XGHETCC	Central city minority market. A 0-1 dummy variable (1 if minority market is in the central city).	Variables used to test for racial price discrimination
XGHETSB	Suburban minority market. A 0-1 dummy variable (1 if minority market is in the suburbs).	Variables used to test for racial price discrimination
XMIXCC	Central city mixed market. A 0-1 dummy variable (1 if racially mixed market is in the central city).	Variables used to test for racial price discrimination
XMIXSUB	Suburban mixed market. A 0-1 dummy variable (1 if racially mixed market is in the suburbs).	Variables used to test for racial price discrimination
XWHITCC	Central city white market. A 0-1 dummy variable (1 if white market is in the central city).	Variables used to test for racial price discrimination
XWHITSB	Suburban white market. A 0-1 dummy variable (1 if white market is in the suburbs).	Variables used to test for racial price discrimination

APPENDIX VII

DESCRIPTION OF PRINCIPAL COMPONENTS ANALYSIS AND LISTING OF ROTATED FACTOR MATRICES

The first section of this appendix describes the application of principal components analysis in the present study. Principal components analysis was used for three sets of data: surface, structure, and window quality; Census tract descriptors; and participant responses from the Baseline Interview aggregated over groups of Census tracts (C* neighborhoods). In the second section the rotated factor matrices for these principal components analyses in Pittsburgh and Phoenix are presented.

ORTHOGONALIZATION OF A MULTICOLLINEAR SET OF VARIABLES

Multiple regression becomes problematic when the predictor variables are highly collinear. The resulting large variance of estimate of the coefficients may make them effectively uninterpretable. When the covariances are nil, on the other hand, the relation of any single predictor variable to any external criterion is independent of any other predictor variable's relation to that criterion and can be interpreted separately and unambiguously.

Uncorrelated predictor variables, may be constructed by transforming the original data. Linear transformations of the vector of predictor variables, "z," which adequately represent the original variation in the predictor set and yet are uncorrelated with each other are the type of transformation most frequently and coveniently used.

PRINCIPAL COMPONENTS ANALYSIS

Let "z" be a vector of standardized predictor variables. There is one such vector for each unit of analysis. Let "y" be a general linear transformation of "z" of the form

$$y = V'z$$

where V is a JxK coefficient matrix that carries the J element variable "z" into the derived K element variable, "y". If the "z" variables are standardized, the dispersion matrix of "y" is given by

$$D_Y = V'RV,$$

where R is the correlation matrix for "z", since the dispersion of a set of standardized variables is a correlation matrix. Thus, the elements of the derived vector, "y", will be uncorrelated if D_Y is diagonal. There are infinitely many transforms, V′, which will diagonalize R; hence, further restrictions must be imposed on the problem. One such restriction is that the variances of the leading elements of "y" be maximized. Harold Hotelling's (1933) derivation of principal components analysis was based on this approach.

The solution is well known. The desired set of coefficient vectors, V, are simply the eigenvectors of the matrix R. A dual solution to the factor analysis problem $Z = V'y$ is also provided by the eigenvectors of R. These solutions are called principal factors.[1] Although conceptually factors are quite different from components the distinction is commonly ignored. In fact some authors, for example Cooley and Lohnes (1971), simply define the term factor as a standardized component. In any event, since the "best" estimate of the factor is given by the normalized components, that practice will be followed in this appendix.

ORTHOGONAL ROTATION TO SIMPLE STRUCTURE

Although principal components analysis defines a unique set of linear transformations of the original set of variables using the maximum

---

[1] One important feature of principal factors is that while the complete set of J factors will exactly reproduce the correlation matrix, R, and thus account for the vector variable, "z", it is possible to retain in a research solution only the first K factors with confidence that they (or orthogonal transformations of them) extract more of the trace of R than any other set of K orthogonal components.

variance criterion as described, the procedure does not usually result in a satisfactory set of such transformations from an interpretive point of view. Since Thurstone (1947) set forth the principles of simple structure, data analysts have been interested in schemes for improving upon the solution offered by principal components by further transforming the component coefficients in ways which help to make the derived variables more interpretable, while preserving the tractability and utility of the components solution.

The retained principal components may be normalized to have unit variance by dividing each component by its standard deviation and forming

$$y = L^{-1}V'z,$$

where L is a diagonal matrix whose elements are the square roots of the eigenvalues of R. Any orthogonal transformation of the normalized principal components will be orthogonal and account for the same variance of "z" as the original set.

The results of such rotations are generally interpreted in terms of the factor loadings, S, which transform "y" into "z". Thus S is defined by

$$S = (L^{-1}V')^{-1} = VL$$

since $V'V = I$. The loadings are the correlations of the factors, "y", with the original variable, "z". The loadings are rotated by postmultiplying S by an orthogonal matrix. Assuming that the substantive content of the original variables is well known, the loadings in the k^{th} column of S help in attributing substantive meaning to the k^{th} component. Variables that have high loadings are used to define the component while variables with low loadings are ignored or attributed little interpretive significance.

What is desired in a rotation is a unique transformation of S such that all the loadings, S_{jk}, approach either zero or unity under the restraint that the sum of the squared loadings for any variable remains constant and the orthogonality of the columns of S remain undisturbed. The reason for this is that very high and very low loadings (correlations) make it easier to attach substantive labels to the components. The varimax rotation is one analytical scheme for achieving this end (Kaiser, 1958).

The varimax method seeks simple structure by maximizing the variance of the squared loadings in each column, that is:

$$V = J \sum_{k=1}^{K} \sum_{j=1}^{J} \left(s_{jk/hj} \right)^4 - \sum_{k=1}^{K} \left(\sum_{j=1}^{J} s^2_{jk/h^2 j} \right)^2 :$$

$$\text{where} \quad h_j^2 = \sum_{k=1}^{K} s_{jk}^2 \quad = \text{ commonalty,}$$

under the restriction that the orthgonality of the components and the sum of squares of the loadings in each row remain undisturbed. What the varimax method does is simplify each column or factor by maximizing column variances of squared loadings. Kaiser feels that this represents an analog of simple structure in the sense of leading to interpretable solution.

In general the complete set of elements in S is not rotated, since some reduction of the original data set is often desired. Applications in this analysis have adopted the common convention of rotating only those columns of S whose corresponding eigenvalue was greater than or equal to unity.

CREATING STANDARDIZED COMPONENT SCORES FOR ORTHOGONALLY ROTATED PRINCIPAL COMPONENTS SOLUTIONS

Let S_1 be the matrix of loadings for the rotated subset of S. It can be readily established that the coefficients for computing components scores for the rotated components are defined by the equation

$$B_1 = S_1 (S_1' S_1)^{-1},$$

where B_1 is the coefficient matrix which when transposed and post-multiplied by "z" yields a vector of standardized component scores, "f," for each unit of analysis (Cooley and Lohnes, 1971, pp. 155-158), or in rotational form

$$f = zB_1'.$$

Each unit of analysis will have a vector of scores corresponding to "f." Each element of "f" has a mean of zero, a standard deviation equal to unity, and nil correlation with the remaining elements.

Table VII-1

PITTSBURGH: SURFACE AND STRUCTURAL RATINGS
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1 XF1STR	FACTOR 2: XF2SUR	FACTOR 3 XF3FSUR	FACTOR 4. XF4WIN	FACTOR 5 XF5FSTR	FACTOR 6 XF6BSUR
LRCLSTR	Ceiling structure-living room	.747	244	080	075	173	.030
LRCLSUR	Ceiling surface-living room	.157	708	.058	.106	.143	.095
LRWLSTR	Wall structure-living room	.785	.222	.125	.114	181	.020
LRWLSUR	Wall surface-living room	193	.662	.187	.128	143	.128
LRFLSTR	Floor structure-living room	.278	.209	.130	.058	.762	.034
LRFLSUR	Floor surface-living room	.108	.166	.785	.109	.146	-.044
LRWCOND	Window condition-living room	095	.230	.088	.747	.093	.007
BACLSTR	Ceiling structure-bathroom	673	.057	076	.065	.146	.442
BACLSUR	Ceiling surface-bathroom	.098	.371	.059	.126	.076	713
BAWLSTR	Wall structure-bathroom	.692	004	.148	.056	.126	.439
BAWLSUR	Wall surface-bathroom	156	.338	.164	152	.091	.675
BAFLSTR	Floor structure-bathroom	.263	.100	.215	.149	534	.390
BAFLSUR	Floor surface-bathroom	.144	.127	.673	.104	.162	.347
BAWCOND	Window condition-bathroom	.063	105	.101	.673	.084	.266
KNCLSTR	Ceiling structure-kitchen	.758	.200	.092	.110	.163	.094
KNCLSUR	Ceiling surface-kitchen	173	.640	.131	.112	.086	.170
KNWLSTR	Wall structure-kitchen	.801	.168	.157	.096	.152	.095
KNWLSUR	Wall surface-kitchen	.143	.626	.220	.167	.028	.289
KNFLSTR	Floor structure-kitchen	.365	.176	.150	.127	.694	.096
KNFLSUR	Floor surface-kitchen	.139	200	.740	.074	.098	.064
KNWCOND	Window condition-kitchen	.118	.162	.123	.770	.057	.067
BRCLSTR	Ceiling structure-bedroom	750	.243	.075	.106	.202	-.020
BRCLSUR	Ceiling surface-bedroom	.210	.653	062	.146	.123	.011
BRWLSTR	Wall structure-bedroom	.791	.178	.112	.098	.187	.006
BRWLSUR	Wall surface-bedroom	188	.638	.154	.232	.111	.119
BRFLSTR	Floor structure-bedroom	.284	.155	180	.135	762	.039
BRFLSUR	Floor surface-bedroom	.139	.116	.776	.140	107	.076
BRWCOND	Window condition-bedroom	.145	.172	.089	.781	.097	.009
Percentage of variance explained		(35.7)	(9.2)	(6.2)	(5.4)	(4.3)	(4.0)

SAMPLE All enrolled households
DATA SOURCE: Initial Housing Evaluation Form

Table VII-2

PITTSBURGH: CENSUS TRACT DESCRIPTORS
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1: XCENF01	FACTOR 2- XCENF02	FACTOR 3: XCENF03	FACTOR 4 XCENF04	FACTOR 5 XCENF05
XCTMDVAL	Median value of housing stock in the tract	.202	.724	.477	-.118	-.065
XCTMDGR	Median value of gross rent in the tract	.178	.809	.325	-.240	.019
XCTOWNOC	Percentage of owner-occupied units in the tract	.903	.016	180	286	-.034
XCTDUAGE	Median age of housing stock in the tract	-.298	-.802	.103	-.202	.068
XCTDUAC3	Percentage of dwelling units with central air-conditioning	-.102	.777	.096	.026	.004
XCTSGFAM	Percentage of single-family detached units in the tract	.880	.114	114	.254	.002
XCTMDSCH	Mean number of years of school of persons 25 and over	304	.603	.556	-.193	-.164
XCTMDINC	Median family income in tract	.468	.510	612	.117	.020
XCTSMHSE	Percentage in same houses since 1965 in the tract	.564	-.452	-.015	.339	-.242
XCTBLCOL	Percentage of blue collar (14 or more years) workers	.233	-.552	-.466	.515	.061
XCTCRWD	Percentage of dwelling units with 1.01 or more persons per room	.051	-.299	-.749	.136	.011
XCTGT4DU	Percentage of dwelling units in tract with 5 or more units	-.760	.416	.120	-.106	-.029
XCTBATH	Percentage of dwelling units with more than one bathroom	.645	.446	.305	.092	.020
XCTMNRMS	Average number of rooms per dwelling unit in the tract	.881	142	.183	-.219	-.101
XCTPBLK	Percentage black in the tract	-.118	-.001	-.720	-.574	-.056
XCTSPAM	Percentage Spanish American in the tract	-.031	-.037	-.096	.006	.965
XCTMDHSZ	Median household size in tract	.884	.142	-.160	.091	-.093
XCTPTMGR	Percentage of persons (14 or more years) who work in professional jobs	-.099	.712	.518	-.211	.078
XCTCLSL	Percentage of persons (14 or more years) who work in clerical jobs	.096	.137	.774	-.020	-.088
XCTWOPLB	Percentage of dwelling units lacking 1 or more plumbing facilities	-.733	-.193	-.256	.371	-.066
XCTWOHT	Percentage of dwelling units lacking adequate heat	-.415	-.112	-.585	.027	.051
XCTSHKIT	Percentage of dwelling units with incomplete or shared kitchen facilities	-.708	-.055	-.272	.302	-.104
Percentage of variance explained		(38.1)	(23.0)	(7.7)	(5.0)	(4.7)

SAMPLE. All enrolled households
DATA SOURCE: 1970 Census of Population

Table VII-3

PHOENIX CENSUS TRACT DESCRIPTORS
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1· XCENF01	FACTOR 2: XCENF02	FACTOR 3 XCENF03
XCTMDVAL	Median value of housing stock in the tract	.909	.051	-.196
XCTMDGR	Median value of gross rent in the tract	.860	.149	-.339
XCTOWNOC	Percentage of owner-occupied units in the tract	.263	.871	- 162
XCTDUAGE	Median age of housing stock in the tract	-.477	-.457	.409
XCTDUAC3	Percentage of dwelling units with central air-conditioning	.905	.478	-.294
XCTSGFAM	Percentage of single-family detached units in the tract	-.206	888	.116
XCTMDSCH	Mean number of years of school of persons 25 and over	863	-.108	-.413
XCTMDINC	Median family income in the tract	.874	.300	- 237
XCTSMHSE	Percentage in same house since 1965 in the tract	-.206	.516	.506
XCTBLCOL	Percentage of blue collar (14 or more years)	-.898	304	-.052
XCTCRWD	Percentage of dwelling units with 1.01 or more persons per room	-.804	.350	.256
XCTGT4DU	Percentage of dwelling units in tract with 5 or more units	402	-.791	.176
XCTBATH	Percentage of dwelling units with more than one bathroom	.785	.508	-.057
XCTMNRMS	Average number of rooms per dwelling unit in the tract	.499	.799	-.192
XCTPBLK	Percentage black in the tract	-.386	.040	.369
XCTPSPAM	Percentage Spanish American in the tract	-.785	.138	.360
XCTMDHSZ	Median household size in the tract	-.200	.864	-.031
XCTPTMGR	Percentage of persons (14 or more years) who work in professional jobs	.935	-.075	-.132
XCTCLSL	Percentage of persons (14 or more years) who work in clerical jobs	.849	-.244	-.280
XCTWOPLB	Percentage of dwelling units lacking 1 or more plumbing facilities	-.370	-.126	.856
XCTWOHT	Percentage of dwelling units lacking adequate heat	-.720	.038	.580
XCTSHKIT	Percentage of dwelling units with incomplete or shared kitchen facilities	-.141	- 264	.813
Percentage of variance explained		(51.6)	(22.2)	(7 2)

SAMPLE: All enrolled households
DATA SOURCE. 1970 Census of Population

Table VII-4

PITTSBURGH MODIFIED CENSUS TRACT DESCRIPTORS[a]
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1: XCTF01	FACTOR 2. XCTF02	FACTOR 3 XCTF03
XCTMDVAL	Median value of housing stock in the tract	.204	.655	.568
XCTMDGR	Median value of gross rent in the tract	.145	.559	.703
XCTOWNOC	Percentage of owner-occupied units in the tract	.950	.085	-.004
XCTDUAGE	Median age of housing stock in the tract	-.285	- 049	-.770
XCTDUAC3	Percentage of dwelling units with central air-conditioning	-.100	.285	.633
XCTSGFAM	Percentage of single-family detached units in the tract	.901	.058	.107
XCTMDSCH	Mean number of years of school of persons 25 and over	305	.716	.446
XCTMDINC	Median family income in tract	.500	.638	.362
XCTSMHSE	Percentage of same houses since 1965 in the tract	.586	-.214	-.379
XCTBLCOL	Percentage of blue collar (14 or more years) workers	.264	-.717	- 385
XCTCRWD	Percentage of dwelling units with 1.01 or more persons per room	012	-.764	-.116
XCTGT4DU	Percentage of dwelling units in tract with 5 or more units	-.747	.226	357
XCTBATH	Percentage of dwelling units with more than one bathroom	.638	.382	332
XCTMNRMS	Average number of rooms per dwelling unit in the tract	.844	.288	.094
XCTMDHSZ	Median household size in the tract	.868	-.166	.235
XCTPTMGR	Percentage of persons (14 or more years) who work in professional jobs	-.109	.706	.554
XCTCLSL	Percentage of persons (14 or more years) who work in clerical jobs	.153	.643	.037
XCTWOPLB	Percentage of dwelling units lacking 1 or more plumbing facilities	-.652	-.347	-.179
XCTWOHT	Percentage of dwelling units lacking adequate heat	-.423	- 480	-.063
XCTSHKIT	Percentage of dwelling units with incomplete or shared kitchen facilities	-.634	-.277	- 077
Percentage of variance explained		(41.2)	(25.2)	(7.0)

SAMPLE: All enrolled households
DATA SOURCE 1970 Census of Population
a. Excludes percentage black and Spanish American in census tract.

Table VII-5

PHOENIX MODIFIED CENSUS TRACT DESCRIPTORS[a]
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1· XCTF01	FACTOR 2· XCTF02	FACTOR 3 XCTF03
XCTMDVAL	Median value of housing stock in the tract	.908	001	- 192
XCTMDGR	Median value of gross rent in the tract	.863	.089	-.344
XCTOWNOC	Percentage of owner-occupied units in the tract	.295	.840	-.198
XCTDUAGE	Median age of housing stock in the tract	-.496	-.389	.372
XCTDUAC3	Percentage of dwelling units with central air-conditioning	.904	- 011	-.306
XCTSGFAM	Percentage of single-family detached units in the tract	-.170	.883	.056
XCTMDSCH	Mean number of years of school of persons 25 and over	.845	- 166	-.418
XCTMDINC	Median family income in the tract	.880	.250	-.256
XCTSMHSE	Percentage in same house since 1965 in the tract	-.215	.467	337
XCTBLCOL	Percentage of blue collar (14 or more years)	-.876	347	-.042
XCTCRWD	Percentage of dwelling units with 1.01 or more persons per room	-.757	.391	.275
XCTGT4DU	Percentage of dwelling units in tract with 5 or more units	.360	-.788	.183
XCTBATH	Percentage of dwelling units with more than one bathroom	.803	.465	-.084
XCTMNRMS	Average number of rooms per dwelling unit in the tract	.535	770	-.236
XCTMDHSZ	Median household size in the tract	-.138	.847	-.020
XCTPTMGR	Percentage of persons (14 or more years) who work in professional jobs	.922	-.126	-.154
XCTCLSL	Percentage of persons (14 or more years) who work in clerical jobs	.825	- 299	-.271
XCTWOPLB	Percentage of dwelling units lacking one or more plumbing facilities	-.354	-.056	.928
XCTWCHT	Percentage of dwelling units lacking adequate heat	-.699	.092	.586
XCTSHKIT	Percentage of dwelling units with incomplete or shared kitchen facilities	-.193	-.209	.733
Percentage of variance explained		(52.0)	(24.2)	(7.8)

SAMPLE All enrolled households
DATA SOURCE· 1970 Census of Population
a. Excludes percentage of black and Spanish American in census tract.

Table VII-6

PITTSBURGH INTERVIEW RESPONSES
AGGREGATED BY C* NEIGHBORHOOD
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1 XCNHF11	FACTOR 2. XCNHF12	FACTOR 3 XCNHF13	FACTOR 4· XCHNF14
XHCNPARK	Quality of parking	-.412	.543	-.311	-.141
XHCNLGHT	Quality of street lighting	632	- 378	.334	-.157
XHCNSHOP	Convenience to grocery shopping	.684	.404	-.104	-.242
XHCNGARB	Quality of garbage collection	.236	.589	-.486	-.178
XHCNFIRE	Quality of fire department services	.465	403	-.095	-.555
XHCNPOL	Quality of police protection	.227	.575	-.370	-.593
XHCNTRAN	Quality of public transportation	.646	.109	.288	.022
XHCNTREE	Quality of landscaping	- 066	376	-.560	-.446
XHCNPRAY	Convenience to places of worship	.565	.445	- 025	-.051
XHCNMED	Quality of medical facilities	.588	098	.288	.394
XHCNAREC	Quality of adult recreation	.822	.167	-.286	-.148
XHCNTREC	Quality of teenage recreation	.867	066	-.272	.022
XHCNPLAY	Quality of children's recreation	644	- 090	-.529	.012
XHCNCARE	Quality of daycare facilities	.132	.145	-.054	867
XHCNELM	Quality of elementary schools	.294	.406	-.308	-.451
XHCNJH	Quality of junior high schools	.169	.902	-.064	-.108
XHCNHSCH	Quality of senior high schools	.174	.899	- 142	-.129
XHCNSTR	Problems with street repair	-.142	-.219	.419	.538
XHCNLOUD	Problems with noise	-.040	.133	.860	.018
XHCNMESS	Litter and trash problems	-.192	-.453	.550	.522
XHCNTRAF	Heavy traffic problems	.214	-.181	.857	.201
XHCNDRUG	Drug problems	-.065	-.639	.505	.358
XHCNCRIM	Crime problems	.099	-.365	.146	.647
XHCNABAN	Abandoned buildings	-.438	-.354	.522	.500
XHCNJUNK	Litter in vacant lots	-.211	-.341	.651	.532
Percentage of variance explained		(41.2)	(16.6)	(7.4)	(6.8)

SAMPLE· All enrolled households
DATA SOURCE. Baseline Interview

Table VII-7

PHOENIX: INTERVIEW RESPONSES
AGGREGATED BY C* NEIGHBORHOOD
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1 XCNHF11	FACTOR 2· XCNHF12	FACTOR 3 XCNHF13	FACTOR 4. XCNHF14	FACTOR 5. XCNHF15	FACTOR 6 XCNHF16
XHCNPARK	Quality of parking	.274	-.064	.076	.849	-.224	-.025
XHCNLGHT	Quality of street lighting	.015	.091	-.290	.062	.752	.094
XHCNSHOP	Convenience to grocery shopping	.571	.116	.249	.721	-.030	.119
XHCNGARB	Quality of garbage collection	.362	.212	.049	.493	.167	.679
XHCNFIRE	Quality of fire department service	.525	.230	.005	.194	.237	.694
XHCNPOL	Quality of police protection	.728	.069	-.134	.081	.001	.070
XHCNTRAN	Quality of public transportation	-.077	-.100	-.854	-.167	.227	.324
XHCNTREE	Quality of landscaping	.653	.303	-.210	.480	.297	-.018
XHCNPRAY	Convenience to places of worship	.516	.264	.596	.240	-.303	.073
XHCNMED	Quality of medical facilities	.746	.296	.083	300	-.122	.115
XHCNAREC	Quality of adult recreation	.357	.877	.068	.266	.094	.042
XHCNTREC	Quality of teenage recreation	.112	.944	.103	-.029	.139	.048
XHCNPLAY	Quality of children's recreation	.089	.868	.307	.039	.111	.007
XHCNCARE	Quality of daycare facilities	.398	.464	.175	.536	.063	.083
XHCNELM	Quality of elementary schools	.336	.329	.510	551	.226	.097
XHCNJH	Quality of junior high schools	-.087	.191	917	.078	-.107	.052
XHCNHSCH	Quality of senior high schools	.669	.337	.282	.498	-.184	-.074
XHCNSTR	Problems with street repair	-.592	-.346	-.508	.336	-.087	.120
XHCNLOUD	Problems with noise	-.353	- 177	-.335	-.276	-.120	.689
XHCNMESS	Litter and trash problems	-.746	-.450	-.111	-.384	015	- 146
XHCNTRAF	Heavy traffic problems	.078	.180	.014	-.145	.867	.015
XHCNDRUG	Drug problems	-.758	.031	-.336	-.251	-.048	-.082
XHCNCRIM	Crime problems	-.731	037	-.255	-.229	-.066	.194
XHCNABAN	Abandoned buildings	-.717	-.151	-.053	- 520	-.109	- 102
XHCNJUNK	Litter in vacant lots	-.746	- 403	.104	-.096	-.224	-.149
Percentage of variance explained		(45.6)	(13 3)	(10 5)	(5 9)	(4.6)	(4.3)

SAMPLE: All enrolled households
DATA SOURCE· Baseline Interview

Table VII-8

PITTSBURGH: INTERVIEW RESPONSES
AGGREGATED BY C* NEIGHBORHOOD
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1: XCNHF21	FACTOR 2: XCNHF22	FACTOR 3: XCNHF23
XHCNBPOL	Poorly rated police protection	.577	.665	.099
XHCNBTRE	Poorly rated landscaping	.764	.281	.295
XHCNBPRK	Poorly rated availability of parking	.448	.082	-.336
XHCNBRCA	Poorly rated adult recreation	.304	.254	.752
XHCNBRCT	Poorly rated teenage recreation	.070	.212	.860
XHCNBRCK	Poorly rated children's recreation	.159	.020	.837
XHCNBELM	Poorly rated elementary schools	.094	.759	.271
XHCNBJH	Poorly rated junior high schools	.223	.888	.186
XHCNBSH	Poorly rated senior high schools	.182	.905	.102
XHCNSTR	Problems with street repair	.639	.240	.197
XHCNLOUD	Problems with noise	.764	-.027	.266
XHCNMESS	Litter and trash problems	.741	.507	.169
XHCNTRAF	Heavy traffic problems	.813	.116	.032
XHCNDRUG	Drug problems	.776	.438	.013
XHCNCRIM	Crime problems	.430	.574	-.161
XHCNABAN	Abandoned buildings	.597	.505	.438
XHCNJUNK	Litter in vacant lots	.777	.416	.266
Percentage of variance explained		(49.9)	(12.0)	(9.7)

SAMPLE: All enrolled households
DATA SOURCE: Baseline Interview

Table VII-9

PHOENIX INTERVIEW RESPONSES
AGGREGATED BY C* NEIGHBORHOOD
VARIMAX ROTATED FACTOR MATRIX

ACRONYM	VARIABLE DESCRIPTION	FACTOR 1 XCNHF21	FACTOR 2 XCNHF22	FACTOR 3 XCNHF23	FACTOR 4 XCNHF24	FACTOR 5 XCNHF25
XHCNBPOL	Poorly rated police protection	.217	.780	− 088	.058	099
XHCNBTRE	Poorly rated landscaping	.432	.509	.234	.396	-.084
XHCNBPRK	Poorly rated availability of parking	-.303	225	.589	.198	.500
XHCNBRCA	Poorly rated adult recreation	.877	.289	.189	.241	-.020
XHCNBRCT	Poorly rated teenage recreation	.903	.070	053	.107	-.283
XHCNBRCK	Poorly rated children's recreation	.935	.142	069	092	069
XHCNBELM	Poorly rated elementary schools	255	028	894	.066	-.095
XHCNBJH	Poorly rated junior high schools	.356	381	.537	.521	.236
XHCNBSH	Poorly rated senior high schools	.608	.569	.310	.073	305
XHCNSTR	Problems with street repair	462	.198	-.184	.665	-.150
XHCNLOUD	Problems with noise	091	.036	.150	.755	.056
XHCNMESS	Litter and trash problems	.188	.745	.311	.427	-.058
XHCNTRAF	Heavy traffic problems	-.038	-.100	-.027	-.019	.899
XHCNDRUG	Drug problems	.002	.531	.099	.710	.051
XHCNCRIM	Crime problems	.125	.373	.401	.631	-.069
XHCNABAN	Abandoned buildings	131	.611	.587	.389	-.015
XHCNJUNK	Litter in vacant lots	.118	.800	.188	.262	-.356
Percentage of variance explained		(46.1)	(14.3)	(9.1)	(6.4)	(6.0)

SAMPLE. All enrolled households
DATA SOURCE: Baseline Interview

REFERENCES

Cooley, William W. and Paul R. Lohnes, Multivariate Data Analysis, New York, John Wiley & Sons, 1971, pp. 155-58.

Hotelling, Harold, "Analysis of a Complex of Statistical Variables into Principal Components," Journal of Educational Psychology, vol. 24, 1933, pp. 417-441, 498-520.

Kaiser, Henry, "The Varimax Criteria for Analytic Rotation in Factor Analysis," Psychometrika, vol. 23, 1958, pp. 187-200.

Thurstone, L. L., Multiple Factor Analysis, Chicago, University of Chicago Press, 1947.

APPENDIX VIII

ASSESSMENT OF COLLINEARITY

Section 2.1 describes the approach used to assess collinearity among the housing attribute variables. Tables VIII-1 and VIII-2 of this appendix list the variance inflation factors--the diagonal elements of the inverse of the variance-covariance matrix of the independent variables (standardized). Also, the determinant of the correlation matrix is listed. The variables and sample are the same as those for Tables 3-2 to 3-5, the final full sample hedonic regressions.

Table VIII-1

ASSESSMENT OF COLLINEARITY. PITTSBURGH

VARIABLE DESCRIPTION		VARIANCE INFLATION FACTORS	
		LINEAR	SEMILOG
Tenure Characteristics	Related to landlord (0,1)	1.098	1.100
	Length of residence (exponential function)	1.093	1.094
	Landlord lives in the building (0,1)	1.105	1.118
	Number of persons per room	1.148	1.149
	Number of landlord contacts for maintenance	1.113	1.178
Dwelling Unit Features	Area per room (natural log)	1.154	1.163
	Total number of rooms (includes kitchen & bath) (natural log)	1.295	1.382
	Building age (years)	1.673	1.688
	Stove and refrigerator provided (0,1)	1.567	1.580
	Inferior or no heat (0,1)	1.295	1.300
	Garage provided (0,1)	1.106	1.108
	Offstreet parking provided (0,1)	1.125	1.126
	Overall evaluator rating (4 point scale)	1.772	1.796
	Dishwasher and/or disposal provided (0,1)	1.125	1.126
	Recent interior painting or papering (0,1)	1.070	1.068
	Many high quality features (0,1)	1.204	1.207
	Poor wall and ceiling surface (factor score)	1.346	1.350
	Poor window condition (factor score)	1 242	1.244
	Poor bathroom wall and ceiling surface (factor score)	1 171	1 173
	High quality kitchen (0,1)	1.204	1.208
	Presence of adequate exits (0,1)	1.123	1.134
	Air-conditioning present (0,1)	1.155	1.162
	Presence of adequate ceiling height (0,1)	1.101	1.104
	Adequate kitchen facilities present (0,1)	1.039	1.069
	Large multifamily structure (0,1)	1.418	1.458
	Working condition of plumbing (5 point scale)	N/A	1.166
	Presence of private yard (0,1)	N/A	1.257
Neighborhood Features	Good recreational facilities and access (factor score)	1.201	1.203
	Traffic and litter problems (factor score)	1.527	1.540
	Problems with crime and public services (factor score)	1.229	1.235
	Census tracts with higher priced units and higher socioeconomic status (factor score)	1.539	1.538
	Nonminority census tracts with higher socio-economic status (factor score)	1.675	1.705
	Blue collar workers and nonminority residents in census tracts (factor score)	1.068	1.073
	High quality block face (0,1)	1.310	1.314
DETERMINANT OF CORRELATION MATRIX		0 0167	0.0112

SAMPLE. All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.

Table VIII-2

ASSESSMENT OF COLLINEARITY PHOENIX

VARIABLE DESCRIPTION		VARIANCE INFLATION FACTORS	
		LINEAR	SEMILOG
Tenure Characteristics	Related to landlord (0,1)	1 046	1.055
	Length of residence (exponential function)	1 222	1.226
	Number of persons per room	1.282	1.286
	Number of landlord contacts for maintenance	1.043	1.045
Dwelling Unit Features	Area per room (natural log)	1.188	1.227
	Total number of rooms (includes kitchen & bath) (natural log)	1.300	1.418
	Building age (years)	2.675	2.697
	Stove or refrigerator provided (0,1)	1.462	1.469
	Central heat present (0,1)	2 477	2.541
	Garage or carport provided (0,1)	1.214	1.215
	Dishwasher and/or disposal provided (0,1)	1.622	1.625
	Recent interior painting or papering (0,1)	1.102	1.104
	Average surface and structural quality (4 point scale)	3.329	3.979
	Adequate light and ventilation (0,1)	1.281	1.284
	Central air-conditioning present (0,1)	2.671	2.683
	Large multifamily structure (0,1)	1.410	1.410
	Plumbing present (0,1)	N/A	1.436
	Inferior or no heat (0,1)	N/A	2.016
	Presence of adequate ceiling height (0,1)	N/A	1.142
Neighborhood Features	Overall neighborhood quality (factor score)	1.878	1.883
	Recreational facilities (factor score)	1.474	1.477
	Access to shopping and parking (factor score)	1.950	1.962
	Census tracts with higher priced units and higher socioeconomic status (factor score)	3.198	3.266
	Owner-occupied single family dwelling units in census tract (factor score)	1.647	1.659
	Poor quality housing in census tract (factor score)	1.445	1.507
	Distance from the Central Business District (miles)	1.446	1.458
	Quality of block face landscaping (4 point scale)	1.150	1.158
DETERMINANT OF CORRELATION MATRIX		0.0012	0.0004

SAMPLE: All enrolled households, excluding those that moved between the Baseline Interview and enrollment, those with extreme values for residuals, and those living in a neighborhood with fewer than five enrolled households.

DATA SOURCES: Baseline Interview, Initial Household Report Form, Housing Evaluation Form, 1970 Census of Population.