A DISCRETE CHOICE MODEL OF HOUSING SELECTION BY LOW-INCOME URBAN RENTERS

by

Margaret S. Murray

Dissertation Submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Environmental Design and Planning

Approved:

Anna Hardman, Chairman

David Conn

Rosemary C. Goss

C. Theodore Koebel

John Levy

November 1992
Blacksburg, Virginia

A DISCRETE CHOICE MODEL OF HOUSING SELECTION BY LOW-INCOME URBAN RENTERS

by

Margaret S. Murray

Committee Chair: Anna Hardman
Urban Affairs and Planning

(ABSTRACT)

The purpose of this study was to (1) develop a statistical model that classifies the housing problems of affordability, quality and crowding as elements in a choice set facing low-income urban renters, (2) identify the demographic, socioeconomic and regional factors which influence housing outcomes, (3) use the resulting model to estimate the probability that an individual household faces a particular combination of housing problems, and (4) consider how the receipt of housing assistance alters those probabilities by addressing the question of targeting assistance.

The study used data from the American Housing Survey of 1989 to estimate the model. These were low-income renters who both lived in an urban area and moved from one residence to another during the prior twelve month period. The mean income level of the final sample was $14,336. Sixty-one percent of the sample had affordability problems, twenty-eight percent had quality problems and six percent had crowding problems.

The theoretical framework for the study is discrete choice analysis based on a random utility function. The conceptual framework included the development of seven, binary, logit models. These models represent a sequence of choices which the low-income renter makes when finding housing. The assumed choice hierarchy was affordability decisions followed by quality decisions and finally crowding decisions.

The affordability and crowding models performed well; however, the quality model was somewhat disappointing. It appears that either quality is not easily modeled using a binary variable or households perceive quality differently than do the policy makers who establish quality guidelines. The models clearly show that affordability problems constitute the biggest hurdle for the low-income renter. Regional location is a significant factor in estimating the probability of having housing problems. Households in the western region of the United States are most likely to have multiple problems. A major contribution of this study is the focus on housing assistance and how receiving assistance alters the probability of low-income households finding basic shelter.

ACKNOWLEDGMENT

I am grateful to many people for helping me in the preparation of this dissertation. I particularly appreciate the help of my advisor, Dr. Anna Hardman. She read many drafts of this manuscript, each time providing challenging comments. Her persistent efforts to meliorate my writing style greatly improved the final product.

Two people are directly responsible for nourishing my interest in housing. Dr. Rosemary Carucci Goss introduced me to housing as a construct of society and Dr. C. Theodore Koebel showed me how competent research contributes to our understanding of housing problems. Dr. Koebel spent many hours helping me design the models used in this research. I am indebted to them both.

Dr. W. David Conn and Dr. John Levy are two other members of my dissertation committee. They have graciously given me their time and suggestions in committee meetings. My sincere thanks to them. This list of committee members must include Dr. Francis Ventre, my first doctoral advisor. His positive attitude helped me get through the early stages of my doctoral program. I would also like to thank the other faculty members in the Department of Urban Affairs and Planning at VPI & SU. Their classes and seminars provided me with a foundation for understanding the role of theory in the production of knowledge.

On a personal note, I want to especially thank my husband, Tom. He has truly been a steadfast companion during this process. His emotional support encouraged me to persevere and without his technical assistance this dissertation would still be on a computer disk.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

This dissertation integrates two distinct literatures on housing; one from a housing policy perspective and one from the perspective of an analyst concerned with the variables which influence housing decisions. The study's basic hypothesis is that low-income, urban, renter households reveal preferences about the cost, quality and size of a housing unit when they make a housing decision. To understand the contribution made by this research, one must consider both the way housing policy literature treats the relationship between affordability, quality and crowding and the way that empirical studies model housing choice.

Descriptive accounts which reveal the proportion of households with multiple housing problems frequently guide housing policy debates. These accounts leave unresolved the question of what chance a particular household has of finding affordable, adequate quality, uncrowded housing. In contrast, this dissertation uses an econometric technique, discrete choice analysis, to model housing problems. This technique permits estimation of the probability that a household selects a particular combination of housing alternatives from a previously identified set of discrete alternatives. The model predicts not only the probability of a housing alternative

but also identifies those socioeconomic and demographic variables which significantly affect the outcome.

This study focuses first on literature related to factors which shape housing policy and subsequently on housing policy itself. It then addresses issues related to modeling housing markets with particular attention to the impact of housing policy on the housing consumption of low-income renters. Finally, the literature review looks at previous models of housing decisions as discrete choices faced by a household.

The theoretical framework for the empirical analysis is a random utility model of consumer choice which characterizes utility as having both an observed component and an unobserved component. The unobserved component is random. That is, it is not constant across households. The presence of an unobserved, random component permits the estimation of choice probabilities. A description of the socioeconomic and demographic characteristics of the households included in the empirical study precedes the empirical analysis. The analysis uses a sample of low-income households from the 1989 American Housing Survey. The sampled households are categorized by the amount of income spent on housing, the quality of their housing unit and if they occupy a unit defined as either crowded or not crowded.

The empirical portion of the paper presents a series of logistic regression models which connect the probability of finding affordable, adequate quality, uncrowded housing to characteristics of the household, the location of the housing unit, whether the unit is single family or multi-family and, in the case of housing quality, the age of the housing unit. These models assume that each household maximizes utility constrained by income and by the ability to access information

about the housing market. The model uses a hierarchical decision structure. First, the affordability model calculates the probability that a household will spend thirty percent or less of income on housing. This is followed by the estimation of quality and crowding models. The affordability model uses the entire data set. Subsequent models use subsets of the data.

Evidence presented in the analysis lends support to the popular assertion that the primary problem facing low-income renters is one of affordability. Estimates of the probability of living in adequate quality housing depend on the prior determination of affordability. The data is divided into two groups, those households spending thirty percent or less on housing and those households spending more than thirty percent on housing. The last models estimated are the crowding models. These models predict the probability that a household occupies an uncrowded unit. Each crowding model is conditioned on earlier choices of affordability and quality.

The analysis combines probabilities calculated from each individual model in order to estimate the joint probability of multiple housing problems. The resulting joint probabilities focus on three selected household types: a household with a married, male head; a household with an unmarried, female head; and a household with an unmarried, elderly, female head. Income levels for these joint probabilities are fixed at both thirty-five percent and eighty percent of regional median income.

A major accomplishment of this paper is to extend the main strategy used by housing policy analysts from one of using aggregate data and simple descriptive statistics to one of using disaggregate data with a focus on the individual decision making unit, the household. By demonstrating the use of discrete choice analysis in conjunction with housing policy variables, this research examines the tradeoffs

between cost, quality and size that a household makes. This method also highlights the complex nature of housing problems and the effect that belonging to particular socioeconomic and demographic groups has on the probability of multiple housing problems. This research provides an efficient way to evaluate the result of targeting housing resources on the housing outcomes of particular population groups.

1.1 Housing Problems and Housing Policy

This research evaluates empirically the applicability of housing deprivation models to residential choices made by low-income urban renters. That information is then used to address questions about who should be served by housing assistance programs. Housing deprivation is defined as a high rent burden, a low quality unit, a crowded unit, or some combination thereof (Newman and Struyk 1983, Frieden and Solomon 1977). It is necessary to identify elements of housing deprivation both because of the complexity of the housing commodity and because this nation has a commitment to help people reach a basic level of consumption.

With the passage of the National Affordable Housing Act in 1990[1], Congress changed the focus of targeted housing assistance programs by increasing the proportion of households in the fifty to eighty percent of median income group eligible for housing assistance (Nelson and Khadduri 1992). Formerly, assistance programs offered preferential treatment for households below fifty percent of median income with substantial rent burdens or very inadequate housing. The systematic evaluation of the housing trade-offs made by low-income households and the targeting of households with the most severe problems is of fundamental importance in establishing housing policy.

[1] Also known as the Cranston-Gonzales Act after its sponsors in congress.

One stated goal of United States housing policy is to improve the housing condition of the poor[2]. A variety of government programs initiated since the original Housing Act of 1937 address the condition, availability and affordability of housing for the poor. Many researchers find these programs to be both inefficient and inequitable. However, Peter Salins (1987, p.8) points out that by "historical or international standards even the housing of our poorest families is not so bad." Yet, he also asks questions related to how bad the housing conditions of the housing disadvantaged poor are, and about who among the poor are most likely to need assistance. The likelihood of multiple housing problems in the non-assisted low-income group relative to that in the group who participate in government programs is an important component of this research.

One aspect of the low-income housing problem that receives little attention is the housing condition of those low-income individuals who have housing (owners or renters), but who do not participate in any type of welfare assistance program. A study comparing characteristics of tenants in the major housing subsidy programs noted that eighty-seven per-cent of the poor remain unserved (Burke, 1982)[3]. Households with incomes under fifty per-cent of the median income were considered poor. This horizontal inequity (unequal treatment of households in the same income group) has long been a concern of economists and those involved in formulating housing policy (Salins, 1987). One expects that a certain degree of luck is involved with admittance into "the club" of those who receive housing assistance. The cards, however, may be stacked against certain groups. This research considers

[2] In this context housing policy serves as social welfare policy (Hayes 1985). Other goals are as community development policy and macroeconomic policy.

[3] Burke's study was done in order to provide detailed data on types of families served by public housing, Section 8, Section 236 and Rent Supplements in 1974, 1979 and 1981.

the housing outcomes of those not admitted to "the club" as well as the housing outcomes[4] of those receiving assistance.

Much of the current research addressing the housing choice dilemma faced by the low-income renter in securing shelter of adequate quality and affordability is directed at reporting patterns of housing problems. This research provides insight into the probability that a particular combination of problems will exist. It also evaluates the effects of government housing policies on the housing choice dilemma by examining revealed preferences of both the assisted and non-assisted groups. The basis for many housing studies is aggregate demand and supply at the market level, not the disaggregate choice process at the household level.

1.2 Theoretical Framework

Rouwendal (1989, p.1) contends that prevailing micro-economic models employing continuous variables are not the most natural way to investigate bundles of goods, such as housing, which are available in a limited number of combinations. He suggests instead the use of discrete choice theory as a general framework for analysis. Discrete choice models permit the analysis of discontinuous variables and of the choice probabilities of individuals. An additional advantage to using this type of modeling is the ability to explicitly incorporate household location into the choice process.

Most discrete choice studies on housing focus on either the tenure decision or the housing type decision (McFadden, 1978; King, 1980; Quigley, 1983; Deurloo et. al., 1988; Börsch-Supan, 1988, 1990). Tenure choice models estimate the probability

[4] Throughout this paper, I use the word "choice" in association with the housing decisions made by low-income urban renters. It should be clear that the use of this word is in keeping with the type of analysis used not because I believe that these households face true choices in the marketplace. Where possible, the word choice is replaced by the word "find".

of buying versus renting and housing type models evaluate the selection of a particular style of unit. Tenure models are not relevant to the analysis of housing decisions made by most low-income households, nor are models of housing types such as single family detached, duplex etc. First, low-income households rarely have the necessary down-payment or the stability of income required to make a purchase decision rather than a rent decision. Second, lack of access to information and markets may prevent low-income households from perceiving the choices available, or the local market may not offer choices between single-family and multi-family housing at the price and quality level needed by the low-income household.

Apgar (1990) argues that the housing condition of low-income renters has worsened during the past two decades due to the deterioration or elimination of much of the low-income rental stock. This effectively alters the dimensions (cost, size) of the choice set. Government subsidies may contribute to a decline in the non-subsidized rental stock, thus reducing the choices available to this group (Roistatcher 1981). Broley, Tate and Walters (1976) addressed the issue of consumer preferences in housing and suggested that low-income renters believe that they have no choice about many aspects of their dwellings. They also state that many FHA and HUD policies (i.e., mandatory contributions of a certain percentage of income to rent) do not enhance opportunities for choice because of requirements placed on recipients.

1.3 Problem Statement

The question of availability of adequate housing for low-income families is an ongoing debate. Since the 1930s housing assistance has been a component of social welfare policies directed at helping people consume a basic level of goods and

services. While there is general agreement on who makes up the low-income population[5], there is little agreement on the availability or adequacy of housing for this group.

Surprisingly, little systematic knowledge exists about the housing choice patterns of the low-income population with respect to the trade-offs between affordability, quality and crowding they face. Also, we do not know how the non-assisted population fares relative to those who receive help from the government. Apgar (1991) contends that poverty-level families face a quality/cost trade-off. To support that contention, he has compiled a series of statistics from the Annual/American Housing Survey of 1974, 1980 and 1987.[6] He describes the quality/cost trade-off as the "tough choice." However, Apgar's analysis does not provide any information about how the individual household makes that choice. This study is an attempt to provide that information.

This research addresses three questions:

1. Can a discrete choice model represent housing choices made by low-income renters with regard to affordability, quality and crowding? Will this methodology provide a means of validating the survey-based estimates of housing problems that dominate the literature?
2. Can this model be used to estimate the joint probability of multiple housing problems in specific demographic groups and in specific regions of the country?

[5] For purposes of this paper the definition of very low-income is households with incomes under 50% of median income. Low-income includes households whose income range is between fifty percent and eighty percent of median income. This definition is consistent with the requirements of HUD Section 8 programs.

[6] This survey was called the Annual Housing Survey until 1984 when the name was changed to American Housing Survey. The change was made because data is no longer collected annually.

3. Will this model improve planning and policy decisions by helping inform planning and policy making bodies about the tradeoffs low-income households make between price, quality and crowding? In particular, can the model provide information about the desirability of targeting scarce housing resources?

1.4 Organization of the Dissertation

The organization of this dissertation is as follows: There are six chapters and an appendix which contains a description of the quality model used in the research. Following this introductory chapter, Chapter 2 provides a review of literature relevant to housing policy and housing choice. Chapter 3 develops the methodology for the study. Chapter 4 describes the data and includes tables of descriptive statistics related to the data. The empirical analysis is presented in Chapter 5, and Chapter 6 contains a summary and conclusions.

CHAPTER 2
LITERATURE REVIEW

The primary focus of this dissertation is a probabilistic model of low-income renters' housing condition. The standard methodology for examining housing outcomes is simple enumeration of the percentage of households with affordability, quality and crowding problems. This dissertation offers an alternative approach based on a random utility model. The purpose of this chapter is to present literature related to housing policy and to the evaluation of housing problems, and finally to examine the literature which "sets the stage" for the empirical model developed in this dissertation.

This review begins with a discussion of housing indicators and standards, the measurement of affordability and quality and the concept of filtering. The next section considers literature related to housing policy, the relation of poverty to housing and the role of econometric studies in the measurement of housing problems. The last part of the chapter describes the methodology used in the empirical portion of this dissertation. It includes a discussion of the theoretical and empirical work which make up the housing choice literature.

2.1 Housing Indicators and Standards

Housing policy specifically aimed at low-income households[1] has typically relied on simple measures or rules of thumb to define housing need. These measures tend to evaluate housing in physical terms not in terms of access to jobs or public services. We define these measures using housing indicators and housing standards. Although location and neighborhood problems are recognized as contributing to housing deprivation, their contribution is not easily evaluated. Therefore, no indicators or standards exist to measure them.

Housing indicators are the measures used to evaluate housing problems. The indicators used to identify housing deprivation (where housing deprivation is defined in terms of affordability, quality and crowding) are a rent to income ratio, determination of physical condition and the presence of essential facilities, and persons per room. These indicators began to influence United States housing policy in the 1930s[2].

In contrast, standards reflect normative judgments about the dimension of the indicator. For example, crowding is measured by the persons per room indictor. The current United States[3] standard is that more than one person per room indicates crowding. The indicator has not changed over time but the standard has. At one time a housing unit was not perceived as crowded until the number of persons per room exceeded two. Similar examples of changing standards exist with regard to

---

[1] Housing policy not aimed at low-income households includes favorable tax treatment of mortgage interest and property taxes.

[2] The quality indicators were developed for use in initial efforts to evaluate housing quality following the depression. They were incorporated into the decennial census and later abandoned. Since 1973, the Annual/American Housing survey has included quality related questions.

[3] Researchers who study third-world countries rightly point out the cultural bias in this measure and note that in most countries the standard would be much higher.

affordability and quality (Baer 1976, Feins and Lane 1981).

Housing deprivation is a popular way of referring to households whose housing conditions fall below current standards. It is typically used by researchers to draw attention to the numbers of households who inhabit housing that is expensive relative to their income, of poor quality or in poor physical condition and/or overcrowded (Nelson and Khadduri 1992, Apgar 1991, Newman and Struyk 1983, Frieden and Solomon 1977). This thesis represents the first attempt to use indicators (affordability, quality and crowding) of poor housing as elements in the choice set confronting low-income households when they are faced with making a housing decision.

2.1.1 Studies of Affordability and Housing Quality

The literature pertaining to determination of affordability and housing quality is reviewed below. The first subsection discusses affordability. The second subsection contains a review of methods associated with measuring quality and of published research related to quality changes. A subsection on filtering, the process of a house moving from one quality level to another, explains the relationship between quality and price.

Affordability. The burden of costly housing is of growing concern. Apgar (1991, p. 23) states that the cost of obtaining housing has superseded inadequate quality as the primary housing problem of the poor. In describing the housing problems facing low-income households, he argues that the affordability problem reflects the loss of low-cost housing as a result of disinvestment and selective upgrading. Apgar uses American Housing Survey data from 1974, 1980 and 1987 to substantiate his claim that more and more low-income households are faced with

housing cost burdens in excess of thirty percent of income. He also claims that any increase in the number of poor households able to secure adequate and affordable housing since 1974 is due to an increase in the number of subsidized households. The data Apgar compiled indicates that the number of low-income, unsubsidized, renter households living in costly units has more than doubled since 1974.

The issue of increasing rent burdens in the face of declining real incomes among the poor is addressed by Schwartz, Ferlauto and Hoffman (1988). They review a number of housing affordability studies and conclude that the housing crisis for low-income tenants is that the poor have less money to pay for apartments which are declining in quality. Neither Schwartz, Ferlauto and Hoffman nor Apgar consider regional differences in housing affordability problems. Also, they do not address the demographic composition of groups with housing problems. The empirical analysis presented in Chapter 5 considers both issues.

Measuring Housing Quality. The methodology associated with measuring quality receives a great deal of attention in the literature on housing. Housing quality has long been a concern of government policy makers. Early attempts to estimate the quality of housing units noted the presence or absence of essential plumbing and whether the unit needed major repairs. The 1940 Census of Population and Housing included an attempt by the government to evaluate housing quality. The estimation methods changed with each subsequent census until in 1970 the effort was dropped completely. The Annual Housing Survey (later the American Housing Survey) is, since 1973, the principal source of housing quality data.

Many researchers today employ an econometric technique, hedonic regression analysis, to evaluate the quality of housing units. This technique

disaggregates the price or rent of a dwelling unit into individual dwelling attributes. The quality attributes of rental property were considered by Wieand (1983). He used AHS data to illustrate the relationship between quality and rent within and among metropolitan areas. However, he did not look at quality indicators alone in constructing the regression equation; also included were market factors. His analysis showed that performance measures (i.e. not the presence of plumbing but the working condition) are important components of quality. Wieand noted that dummy variables indicating whether or not the unit was subsidized and whether the resident had moved were important. The addition of those variables complicates the analysis. It is not clear that residential mobility is related to housing quality.

The difficulty in finding an objective measure of housing quality is a problem addressed by many researchers. A study by Clemmer and Simonson (1983) examined the trends in substandard housing with respect to quality changes. The authors cite the need to link past to present in establishing a method of measuring housing quality. They defend the HUD definition as "state of the art" and support its use in measuring year-to-year qualitative changes (p. 461). The HUD definition of physically deficient housing or some modification of it is used in a variety of other studies (Newman and Schnare 1988, Newman and Struyk 1983, Frieden and Solomon 1977, Goedert and Goodman 1977).

A recent paper by Spain (1990) related residential mobility to housing quality. She was particularly interested in the role residential mobility played in improving housing quality for female householders. Rather than develop a measure of housing quality, Spain relied on four indicators. The indicators were crowding, market value, gross rent, and tenure. She calculated four regression equations, one for each quality (indicator) variable. The independent variables included mobility

and various demographic characteristics. Spain provided little justification for including those particular items as measures of quality. Since her study lacked a single dependent variable, her conclusions about the relationship between quality and mobility are mixed. That is, she did not find a consistent positive relationship between mobility and quality in all four regressions.

Another method of identifying residential quality appeared in an article by Blank and Rosen (1990). The focus of the study was changes in housing conditions of poor families. A difficulty with studying housing quality is, as the authors point out, the lack of adequate information. However, they proceed to identify housing attributes as characteristics related to the composition of the household[4]. This approach seems to carry with it some built-in problems. The assumption the authors make relating quality to household composition has no clear relation to previous studies of housing quality or to housing quality measures used in the AHS. Nonetheless, the authors conclude that there has been a decrease in "housing independence" (more multiple-family-unit households) during the eleven-year period of study. They do add that this does not necessarily imply that housing conditions have deteriorated (p. 29). Blank and Rosen concluded that real income for urban poor families has fallen during the 1980s, that government housing has declined in availability and that the result is that the housing condition of the poor has been adversely affected (p. 27).

---

[4] The attributes that Blank and Rosen use in this study are: 1. the number of family units per household; 2. the number of persons per household; 3. the percent of family unit heads who are also household heads; 4. the percent of family unit heads who live in households in which the household head is a parent of the family unit head; and 5. the percent of households that own their dwelling.

2.1.2 Filtering

Filtering is credited with improving quality. Houses may filter down to lower income groups when new housing is constructed at a higher quality level. As people move to the new housing, the demand for the type of housing they have vacated decreases, the price falls and the house is occupied by a lower income family[5]. Two issues are important here. One concerns the quality of the older unit and the second concerns the price. The relationship between quality and price has been debated since filtering was first described by Ratclif in 1949. Those who argue that the quality of a unit must decline in order for the price to decline (Lowry 1960) emphasize the fact that an owner or landlord will defer maintenance until the quality of the unit reflects the attainable rent. More recently, the view that declines in quality lag behind rent decreases due to the durability of housing and the inability of landlords to disinvest quickly has become accepted (Braid 1986). The former view suggests that filtering will not improve the overall quality of housing while the latter view says that it will.

Weicher and Thibodeau (1988) developed an empirical model to test for the occurrence of filtering in the 1960s and 1970s. Their model employed data from both the Annual Housing Survey (AHS) and the decennial Censuses. The study was designed to look at the change in occupancy of low-quality housing relative to private housing construction. Their conclusions provide support for the argument that filtering results in lower prices for better quality housing and a reduction of substandard housing. Weicher and Thibodeau also looked for the impact of government subsidies on the incidence of substandard housing. They found that

---

[5] The opposite condition or filtering up also happens particularly in areas of cities where gentrification is occurring.

neither moderate income programs nor subsidized production are effective in reducing the incidence of substandard housing.

The concept of filtering is not confined to housing units. Baer and Williamson (1988) note that many researchers view filtering as a household process related to life cycle variables. This view of filtering emphasizes the movement of households through stages in the life cycle (marriage, birth and rearing of children, old age) and the relationship between these stages and housing choices. Following an extensive review of the literature related to filtering, Baer and Williamson (1988) propose a conceptual framework for unifying alternative dimensions. The framework illustrates the links between: "command over resources, qualities of households and housing units, and location in social and geographic space (p. 141)."

The issue of affordability and the impact of housing quality on household choice are both aspects of the analysis presented in Chapter 5. Filtering emphasizes a change in either the housing unit or a household's circumstance. The household choice model provides a basis for further research directed at evaluating this change. The next section presents a brief history of federal housing policy. It contains a summary of the direction that housing policy has followed with respect to low-income renters since the Depression. Chapter 6 presents suggestions regarding housing policy in light of the housing choices made by households receiving some type of assistance. The material in this section is included in order to provide a background for that discussion.

2.2 Federal Rental Housing Subsidy Programs

The federal government's involvement in rental housing subsidy programs began in the Depression. Indeed, the idea of a rent-to-income ratio probably grew

out of a Housing Act of 1937 requirement that, for most low-income households, tenant income could not surpass five times the rent paid. Another attempt to tie rent paid to income earned was the 1969 Brooke amendment to the Housing Act of 1937. This amendment limited rents in public housing to twenty-five percent of tenant income. A subsequent amendment passed in 1981 raised the ratio to thirty percent.

Prior to 1974, the bulk of federal subsidies to low-income households were tied to increases in the supply of housing. Programs to construct low-income public housing[6] provided loans to local public housing authorities to finance construction. A variety of other programs appeared during the forty year period from the depression until the early nineteen-seventies (see Hays 1985, Jacobs et al 1982). Most of these programs encouraged the construction of housing units or the rehabilitation of existing units. Some programs were tied to the elimination of slums.

In 1974 the passage of the Housing and Community Development Act changed the emphasis from supply to a combination of supply oriented and demand oriented programs. A major contribution of this act was the Section 8 housing subsidy program. This program provides funds to bridge the gap between the market rental of a unit and a fixed percentage of the tenant's income. The amount of the subsidy is also tied to both family size and area medium income.

The Housing Act of 1983 virtually eliminated supply-related housing subsidies and significantly reduced resources available for increases in demand-side subsidies. In most instances, the demand-side subsidies restrict a participant's choice. Under the Section 8 program, the unit must meet certain quality standards,

[6] Other federal programs were aimed at financing homes for middle and upper income families.

it must rent for no more than the area's Fair Market Rent (FMR), and the participants are required to contribute thirty percent (formerly twenty-five percent) of their income to rent. The voucher program, begun during the Reagan administration, gives households more flexibility. In the voucher program, the subsidy is determined by calculating the difference between fair market rent and thirty percent of adjusted income. This system more closely resembles an income transfer, because households finding housing for less than FMR can keep the difference. Alternatively, a household can choose to allocate more than thirty percent of income to rent (Hays 1985).

Housing policy in the 1990s turned a corner with the passage of the National Affordable Housing Act of 1990. This act explicitly recognizes that states and local communities have a responsibility for providing affordable housing. Before tapping in to available federal housing programs, states or localities must develop a Comprehensive Housing Affordability Strategy (CHAS). This strategy requires assessment of both the households needing assistance and the local housing inventory. With these assessments in hand, local housing jurisdictions must set goals for the numbers and types of households needing assistance (Nelson 1992).

2.2.1 Welfare Related Rental Housing Subsidy Programs

As Newman and Schnare (1988, p. 8) point out, the government approach to housing assistance is really "two-pronged." In addition to the programs specifically tied to a housing stream, there are programs tied to the welfare system stream. While the housing stream primarily flows from federal sources, the welfare stream flows from both federal and state sources. Certain types of assistance programs alter housing consumption. Murray (1980, p. 33), in a study of public housing

applicants, found that a housing subsidy increased housing consumption by a significant amount over the level that would have occurred with an equivalent cash transfer. It is generally agreed among economists that the income elasticity of demand among low-income renters is low (Polinski 1977, Mayo 1981). This indicates that some part of a demand-based subsidy will be treated as an income transfer. Price elasticity, the percentage change in quantity of housing service demanded that results from a percentage change in price, is also relevant. Weicher (1979, p. 485) argues that a large part of any subsidy is spent to do what people would do anyway. That is, to support their current housing consumption, not to increase it.

The requirement that a family spend a minimum fraction of income on rent sets an upper limit on the amount available for other, non-housing, goods. Individuals with a weak preference for housing relative to other goods are thus prohibited from exercising that taste if they participate in housing subsidy programs. Economists are quick to point out that, according to the principle of consumer sovereignty, individuals should be free to select a bundle of goods that best satisfy their wants (Stiglitz 1988, p. 355). To the degree that assistance programs act as income transfers, this principle is upheld. Economists also argue that pure income transfers are a more efficient method of redistributing wealth than are housing related programs. Conversely, housing advocates argue that the government does have an interest in maintaining minimal consumption levels. The health and safety of those household members, particularly children, is a necessary component of public policy.

Housing assistance programs are very unequally distributed across the low-income population. The issue of horizontal equity or lack thereof is central to this

research because it affects the probability that one household will obtain assistance in preference to an identical, equally deserving household. Weicher (1979, p. 493) states that "all of the programs are inequitable - they provide a large benefit to a small number of households participating, but nothing to a very large number who are equally eligible." Housing has never been viewed as an open-ended entitlement as has been, for example, food stamps or medicaid (Salins 1987, p. 9). Also, as the value of housing assistance increases, so does the disparity between those who do and do not receive it. The Omnibus Budget Reconciliation Act of 1981 targeted housing assistance to households with incomes below fifty percent of area median. Congress identified preferential groups as those who spend more than fifty percent of income on housing or those who live in very inadequate housing, or who are involuntarily displaced. There is no evidence that benefits are targeted to those least able to secure adequate housing either for reasons of race or sex (Spain 1990).

The targeting of housing assistance is the focus of an article by Nelson and Khadduri (1992). They are concerned that the National Affordable Housing Act of 1990 (the Cranston-Gonzalez Act) reduces targeting. Nelson and Khadduri illustrate the continued need for targeting through a detailed examination of the 1989 AHS data. The worst housing problems are largely confined to households falling below fifty percent of median. These "worst-case" households contain a large proportion of families with children. The analysis in Chapter 5 of this research takes up the issue of targeting and uses the empirical results of this study to show the probability that particular groups will find standard housing with and without housing assistance.

Policies directed at either increasing the supply of housing or at supporting the demand for housing express a normative choice between housing and other

societal needs (education, health care, food). Decisions to allocate funds to housing are made for a variety of reasons, not only because these funds improve the housing condition of recipients. The political expediency of supporting programs aimed at increasing consumption of housing cannot be denied. Tying assistance to quality standards reflects concern for externalities associated with bad housing, and the taxpayers' interest in assurance that the money is not being spent in less desirable ways. Public housing is tangible evidence that the money is being spent on housing (Eekhoff 1987, Aaron 1972).

The arguments for housing subsidies and public housing in preference to a direct cash grant run from market failure on the one hand, to health and safety on the other. The more liberal proponents of housing assistance usually maintain that market mechanisms are not meeting the needs of low-income households and government intervention is necessary to insure that all individuals are adequately housed. Health and safety arguments revolve around a concern for the well-being of families, who are perceived to undervalue housing (Olsen 1982). Likewise, some housing analysts do not believe that direct cash grants are an effective method to use to alleviate housing problems. Their main contention is that cash will only serve to drive up rents, not to increase supply. The Experimental Housing Allowance Program (EHAP) carried out in the 1970s refutes that contention (Mills and Sullivan 1981). No discernible market rent effects were found. Mills and Sullivan do point out that, for a given supply elasticity, if recipients are heavily concentrated in a submarket, then it is more likely that price increases will occur (p. 276).

The next section summarizes four studies which identify factors that relate housing problems to poverty. Of particular note is that none of these studies compare the housing situation of those low-income households who do not receive

assistance to those households receiving some type of assistance. Only when these comparisons are made can questions about the housing problems of all low-income renters be adequately addressed. The analysis in Chapter 5 makes such comparisons.

2.2.2 Empirical Studies on Poverty and Housing

Among the empirical studies which attempt to relate housing deficiencies to poverty is one by Newman and Struyk (1983). Using data from both the Panel Study of Income Dynamics (PSID) and the Annual Housing Survey (AHS), they constructed measures of permanent poverty and housing deprivation. The PSID data was used to predict the probability of permanent poverty and the AHS data was used to measure housing conditions. Newman and Struyk were concerned that housing allocation systems do not differentiate between those who are permanently poor (and thus have little chance of improving their housing condition) and those who are temporarily poor. They concluded that the permanently poor were twice as likely to live in units failing HUD's basic quality definition. The authors argue for targeting allocations of scarce housing resources to those who are permanently poor. Their study did not separate families who receive housing assistance from the entire sample. They did, however, compare differences between permanently poor and temporarily poor owners versus renters, urban versus rural residents, and blacks versus non-blacks.

Newman and Schnare (1988) conducted a comprehensive study of the interaction between welfare programs and the housing situation of low-income households. This study examined the impact of the housing portion of programs like Aid to Families with Dependent Children (AFDC), Supplemental Security Income

(SSI), General Assistance (GA) and HUD programs on the recipient's ability to find adequate housing. Newman and Schnare find that many inequities exist between the programs and that there are many regional differences in the level of support. They do not address the housing situation of low-income households not receiving support. The authors plan to follow this analysis with a report on policy implications of a restructured housing program.

Hammond (1987) used data from the Annual Housing Survey of 1977 to evaluate an intertemporal model of the housing consumption of subsidized households. She found that subsidy programs increased housing consumption over what it would be without the programs. She also found that for some groups (not public housing recipients), benefits increase as income increases. This indicates that programs are not well-focused.

Excessive rent burden and neighborhood inadequacy are addressed by Frieden and Solomon (1977) in a study dealing with the nation's housing during the decade from 1965 to 1975. They measured housing deprivation in different groups of the population and found that, while quality was still a minor problem, the major problem in housing was affordability and neighborhood quality. The definition used in their study is used in this research. However unlike this study, they did not separate assisted and non-assisted low-income renters from the rest of the renter population.

2.2.3 Studies of Housing Markets and Policy

Neoclassical economic theory provides a useful basis for understanding housing markets and the economic implications of policy. This subsection first summarizes the relationships between housing markets and housing policy and

concludes with a discussion of the most recent trends in federal housing policy.

In an early theoretical paper, Sweeney (1974) constructed a model of the urban rental housing market and applied it to selected classes of housing market subsidy programs. A central theme of his model was the quality distribution of housing units. Two main conclusions reached by Sweeney were first, that it is impossible to achieve a simultaneous reduction in rent levels and reduction in the number of low quality units without a demand-side subsidy; and second, that a demand-side subsidy program will adversely affect families who are not subsidized. The model developed by Sweeney also indicated that supply-side subsidies would not reduce the rental prices in the quality ranges desired.

The importance of understanding supply and demand relationships as the key to understanding the welfare implications of subsidy programs is stressed by Hanushek and Quigley (1982). Using data from the Housing Allowance Demand Experiment (HADE)[7], they developed a dynamic model of housing adjustment. Their analysis demonstrated that identical policies may have different results in different housing markets, that adjustment to increasing incomes or decreasing rents may take a long time to produce desired changes in housing consumption, and that the resulting changes depend on the elasticities of housing supply as well as demand.

Another issue is whether public housing and supply oriented government programs reduce the amount of housing available to low-income households or whether these programs increase it. Many economists argue that supply oriented programs substitute for market provision and therefore no increase in total quantity

[7] The HADE (Housing Allowance Demand Experiment) program was conducted in Pittsburgh and Phoenix with low-income renters. The sample was divided into treatment groups who received rent rebates and control groups, those who did not receive rent rebates. Data was collected at three points in time, at the beginning, at t = 1 year, and at t = 2 years.

occurs (Muth 1975). Vitaliano (1983) takes a contradictory view. He developed a model to test if the supply of private rental housing is reduced by the construction of public housing. His model considered the elasticity of supply of low-rent units in the private market and the degree to which public units substituted for private units. The conclusion reached was that (assuming an inelastic supply of low income units) public housing does not reduce the amount of private housing. The author points out that this conclusion may be due to some theoretical or empirical shortcoming. There is also reason to believe that the supply is relatively elastic in the long run (Arnott 1987).

Knowledge of the income and price elasticities of housing demand provided by economists serve as an aid to policymakers. Elasticity measures the percentage change in quantity demanded due to either a unit change in income (income elasticity of demand) or a unit change in price (price elasticity of demand). Mayo (1981) reviewed a number of studies related to the estimation of housing demand. He stressed the importance of including demographic variables in estimation and noted that no single elasticity exists. Elasticities vary "with the level of both relative prices and income, with demographic variables, and with time (p. 113)." This suggests the necessity of estimating separate elasticities for low-income households. Elasticities estimated by Hanushek and Quigley (1982) for low income renters indicate that neither a price nor an income subsidy results in an increase in housing expenditure. The main effect is to reduce rent burden, not improve housing situation. Additional evidence for this view is provided by Venti and Wise (1984). In a study of the relationship between moving and housing expenditures, they found very low income elasticities and indicated that families, where their current housing expenditure was below that minimum, would be little helped by programs that

required a minimum expenditure on housing.

Despite the abundance of research on housing demand elasticities for low-income households and the resulting policy implications, housing policy has not always followed the indicated path. Estimating the probability that particular households have a combination of housing problems and how the receipt of housing assistance alters that probability is an attempt to provide another way of indicating a path for housing policy decisions. The remainder of this chapter describes the methodological approach used in the empirical analysis.

2.3 Methodology

By focusing on the tradeoffs low-income households make between affordability, quality and crowding and the probability associated with particular combinations of problems instead of the percentage of households falling into each category, it is possible to evaluate the choices low-income renters must make when locating housing. The methodological approach described in this section provides the needed structure for evaluating those tradeoffs.

2.3.1 Studies of Housing Choice

The theoretical basis for this research rests on the application of random utility theory to individual spatial choice behavior implemented in a discrete choice framework. A number of housing researchers use the random utility model[8] in discrete choice situations (see Van Lierop 1986, Chapter 9). The random utility model connects individual choice behavior to spatial situations. Including the spatial

[8] Random utility was first described by a psychologist, Thurston, in 1927. In the 1970s it was adopted by transportation researchers to explain travel mode choices. Later, housing analysts applied it to housing choices.

dimension bridges the gap between social theory and urban and regional planning (Karlqvist 1978). That is, people's behavior and well being is shaped by elements which have spatial characteristics. Mode of transportation or housing location, for example, are manifestations of planning decisions which affect behavior.

2.3.2 Random Utility

Random utility theory characterizes utility as composed of a deterministic component and of a stochastic or random component. The deterministic component is directly observable. The random component occurs either because the analyst does not observe all of the variables which influence utility, or because individual behavior is frequently random. That is, an individual does not always select the same alternative when repeatedly presented with identical choice sets (Hensher and Johnson, 1981).

Random utility theory lends itself to modeling housing selection where choice sets are described by discontinuous or discrete alternatives (for example own versus rent). By contrast, traditional consumer theory from economics views utility as a continuous function described by quantities of a commodity. The consumer selects a preferred bundle of commodities limited only by available income. A representative consumer is permitted fractional levels of consumption (Ben-Akiva and Lerman 1985).

Applying traditional utility models to housing selection is likely to produce biased results. Consumers do not simply adjust expenditures in order to consume various amounts of housing but substitute between discrete choices (Börsch-Supan and Pitkin 1988). Discrete choice modeling involves a synthesis of micro-economic consumer theory with psychological theories of choice and statistical analysis

(Wrigley and Longley 1984).

2.3.3 Choice Set Determination

Fisher and Nijkamp (1987, p. 6) point out that "Many important decisions an individual is facing in his (sic) life involve choices from a constrained set of alternatives." The low-income population is particularly constrained in housing alternatives. Housing choices for the low-income household are not always as numerous or well defined as are those for a household with a higher income level. That is, tenure choices, choices between a number of different housing configurations or choices between locations relative to work are not realistic options for much of the low-income population.

McFadden (1974, 1978) was one of the earliest researchers to apply random utility maximization to housing selection using a discrete choice model. His methodology is widely cited by other researchers. McFadden notes that each dwelling consists of a set of attributes among which choice is made. Since no restrictions are placed on the precise attributes involved, researchers select various combinations of attributes for inclusion in the choice set. The research questions asked and the availability of data guide selection of attributes.

In describing methods for modeling residential selection, McFadden (1978) establishes a basis for reducing the size of the choice set (from a possibly infinite size) to a representative set of alternatives without loss of validity. His method involves sampling from the full set of housing alternatives. This is an obvious advantage in residential modeling, where theoretically no two houses are alike due to the unique location enjoyed by each house. In almost all cases, the choice set is truncated in order to simplify the model. The dependent variable in those models is

typically either the type of dwelling unit (single-family, multi-family), or the type of tenure (own or rent), or some combination thereof. The dwelling unit characteristics (number of bedrooms, number of bathrooms, etc.), socioeconomic and demographic variables are the independent variables.

2.3.4 Spatial Issues

Van Lierop (1986, p. 12) describes spatial interactions as the fuel which makes society run. He considers spatial interaction planning as: "the scientific analysis of alternative processes of policy measures designed to achieve particular links between (actors in) nodes in geographic space." These links are related to economic and social goals and, to a lesser degree, environmental or resource management goals. A direct connection between spatial interaction planning and housing policy is easily made. Housing markets are classified in spatial dimensions (neighborhoods, cities, regions), by dwelling type, by tenure, by public versus private housing, and by price. Households make choices related to each of the above and each are in turn influenced by policy decisions. Also, a household's location defines access to public and private services. It is important to "look at the interrelations between housing, location, spatial structures and social practices" (Kearns 1990, p. 123).

Spatial interaction models of housing as described by Van Lierop (1986) generally follow one of three tracks; residential mobility models, housing choice models or relocation models. Residential mobility models focus on housing dissatisfaction due to changes in the socioeconomic characteristics of the household (Goodman 1990). Housing choice models extend the theory of housing demand to include the combined choice of location and housing type (Quigley 1976, Kain and

Apgar 1977, Lerman 1979). And lastly, relocation models attempt to integrate the mobility and choice models into a dynamic framework of the decision to move and settle elsewhere (Van Lierop and Rima 1982, Onaka and Clark 1983, Boehm, 1979). This research follows a housing choice model framework. The households included in the analysis are those who moved recently, not those in the process of searching and moving. By confining the analysis to recent movers, the housing choices made should reflect current market and policy conditions. Additionally, the housing of recent movers may reflect an equilibrium housing situation[9]. Arguably, low-income renters may never reach an equilibrium. Many move because they loose jobs or to find lower cost units not necessarily to improve their housing condition.

2.3.5 Statistical Models

Housing choice problems are frequently evaluated in a probabilistic choice context. Statistical models which forecast the probability that a decision maker selects a particular alternative have been widely used since the 1970s. These models permit the dependent variable to assume discrete values. In the simplest case, the dependent variable is a binary categorical variable (0, 1). More complicated cases involve polychotomous variables or variables which are classified into many categories.

Two models which receive the most attention in evaluating housing choices are the logit model and the probit model. These models can accommodate either a binary or multinomial response (dependent) variable. The primary difference between these models is related to assumptions about the error terms. In the probit

[9] Due to the lower transaction costs associated with moving from one rental dwelling to another, it is more likely that renter households consume an equilibrium level of housing when they move than do owner households.

model, the error terms are distributed normally with a mean of zero and a variance, σ^2. The logit model assumes that the error terms are distributed logistically (Madalla 1983). The coefficients produced by the models are of similar magnitude although estimated on different scales. The logit model is computationally more convenient and it has dominated the housing choice literature.

2.3.6 Housing Market Analysis Using Logit Models

A number of researchers (Börsch-Supan 1988, 1990, Ellickson 1981, King 1980, Quigley 1976, 1983, Kain and Apgar 1977) estimate discrete models of tenure choice or of housing and neighborhood choice. Quigley (1983) reviewed a group of studies that apply the multinomial logit model to housing choice. Researchers variously modeled the consumer's problem as a choice between anywhere from nine housing types to selecting from fifty housing types.

Using random utility theory, Quigley (1976) applied an extended logit model to housing choice and concluded with an empirical example. The model Quigley presented in his paper, the extended logit model, is a joint model of housing choice. In a joint model, the dependent variable is composed of a marginal and conditional probability statement. In Quigley's example, the dependent variable represents the joint probability of selecting both a neighborhood and a dwelling unit.

In a recent article, Börsch-Supan (1990) modeled the tenure choice problem using panel data. A variant of the multinomial logit model was used to perform the statistical analysis. His analysis was confined to a sequence of tenure type and dwelling type combinations. Börsch-Supan found evidence that cross-sectional studies of tenure choice understate the importance of demographic data relative to studies done using panel data.

2.3.7 Sequential Models

Sequential models of housing choice are a means of decomposing the choice process into a series of related decisions. This type of modeling may more closely reflect the actual way that housing decisions are made. That is, the first decision, to move, is followed by decisions regarding location, dwelling type and dwelling unit (Clark and Onaka 1985).

A type of sequential modeling used by Boehm (1979) describes the housing choice problem as a hierarchy of choices related to mobility, tenure, size and quality. He modeled the mobility and tenure decisions as a joint decision. The size and quality decisions were modeled as separate decisions which followed the mobility/tenure decision. This method requires subsetting the data at each level of the hierarchy. Each level of the hierarchy involved a binary decision, move and buy versus move and rent, large versus small, high quality versus low quality. The final model resulted in the estimation of eight different housing combinations (e. g. housing choice one was to move and buy a large high quality house). Boehm calculated the joint choice probabilities for each combination and discussed the relationship between the dependent variables and the independent variables.

Boehm argues that this type of model offers additional insight into the results of government policies on the housing choices of families. He suggests that, once the coefficients of a housing choice system are determined, the effect of housing programs on housing choice could be decided with greater knowledge (p. 143). Ben-Akiva and Lerman (1985) support using a hierarchical model. They propose postulating reasonable choice hierarchies and empirically testing them against each other (p. 55).

The sequential method of modeling a related series of decisions becomes

computationally difficult as the number of choices at each level increases and as the number of levels increases. A form of hierarchical modeling which minimizes the computational burden is the nested logit. In contrast to a hierarchical model that combines marginal and conditional probabilities using subsets of the data, this type of hierarchical model involves a sequential recursive decision structure using the entire data set at each level. The nested logit relies on the estimation of an inclusive value (the log of the denominator of the conditional probability) at each level of the sequence. The estimated parameter of the inclusive value enters the next level as an explanatory variable and serves as a linking mechanism between levels (Hensher 1986).

Clark and Onaka (1985) used a nested multinomial logit model to estimate the joint choice of mobility, neighborhood and dwelling type. The procedure used in a nested multinomial logit model requires estimating submodels in reverse order of the postulated sequence. That is, the submodel for dwelling type is estimated first followed by the submodel for neighborhood and finally the mobility decision submodel. The coefficients estimated at the lower levels are used to calculate the inclusive value for the next submodel. By using a nested multinomial approach, Clark and Onaka were able to estimate dwelling choice from among six types and neighborhood from among nine neighborhood types. The mobility submodel was a binomial logit since the decision is simply move or stay.

Similarly, Börsch-Supan and Pitkin (1988) modeled the housing consumption decision as the joint choice between headship, single or multi-family house, tenure and size using a nested logit model. They estimated a number of alternative model structures and also compared a hierarchical approach to a multinomial approach. Their results confirm the value of a hierarchical approach as a modeling tool.

The nested multinomial logit has the advantage of handling a great number of choices at each level of a sequential decision process[10]. The entire model can be estimated in one step, but most researchers use a sequential method. One problem with the sequential method is that the standard errors of the coefficients of the inclusive values are biased. These errors are then passed through to successive levels of the sequence (McFadden 1981). The nested multinomial logit model differs from the model used in this research in that, while it uses a sequential decision structure, the decision itself is not sequential. That is, a household makes a joint choice not one choice at a time. Another difference is that the nested procedure does not require partitioning the data set as each level of the hierarchy is estimated.

The hierarchical model used in this dissertation does not have the error problem mentioned since the separate models of marginal and conditional probabilities are not linked (i.e. the model is not recursive). The choices are made sequentially not jointly. This type of modeling requires a large data base since the sample size is reduced every time an additional probability equation is estimated.

2.4 Conclusion

The literature cited in this review suggests that ways of understanding the housing problem of low-income households have not yet provided policymakers with a comprehensive solution. If housing policy is to respond to the housing needs of low-income renters, the housing choices made by those households must be explored. Accordingly, the following chapters present a method for examining those housing choices.

[10] If a hierarchical model of combined conditional and marginal probabilities was used in the Clark and Onaka study, it would have required estimation of 108 (6x9x2) combinations.

CHAPTER 3
METHODOLOGY

This chapter develops a detailed model of housing choice based on a random utility function. Corresponding derived probabilistic choice functions estimate the probability that a low-income household will select one housing alternative relative to another. The model originated from the probabilistic choice models which first appeared in mathematical psychology (Thurston 1927, Luce 1959). The first section, Section 3.1, identifies the reasons for using a qualitative model to model housing deprivation. Section 3.2 introduces the probabilistic choice system. Section 3.3 presents an overview of random utility theory underlying the model. A description of the discrete choice model is in Section 3.4, Section 3.5 presents the choice set used in the dissertation, and Section 3.6 describes the method for estimating the probable housing choice[1].

3.1 Modeling Housing Deprivation

Housing deprivation is almost always explained using aggregate data to

[1] Much of the theory which follows is based on the work of Hensher and Johnson (1981), McFadden (1981), Anas (1982) and Ben-Akiva and Lerman (1985).

identify the numbers or proportion of households facing one or more housing problems (Nelson and Khadduri 1992, Apgar 1990, Newman and Schnare 1988, Frieden and Solomon 1977). The need to develop a more comprehensive approach which recognizes the interdependence of housing problems while maintaining a focus on the individual household led to the selection of qualitative choice analysis.

The use of qualitative choice analysis to analyze housing problems in the housing choices made by low-income urban renters presents some difficulty due to the fact that the dependent variables included in the model are arguably continuous, not discrete. However, assigning those variables to categories and thus rendering them discrete is reasonable because the fundamental questions which policymakers pose about housing conditions are categorical. Those questions relate to classifying households as living in affordable, adequate quality, uncrowded housing not measuring how affordable or how adequate or how uncrowded.

Qualitative choice situations determine the probability that a decisionmaker (in this case a household) selects a particular alternative. In qualitative choice analysis, the decisionmaker faces making a selection among a finite set of alternatives. Additionally, the alternatives are assumed to be exhaustive and mutually exclusive (Train, 1986). While utility theory provides the basis for the derivation of qualitative choice models, there is considerable leeway in determining the appropriate form of model. The specific model used depends both on the availability of data and the particular policy questions addressed.

3.2 The Probabilistic Choice System

In general, a model of a probabilistic choice system sets out a set of alternatives, identifies the attributes of those alternatives, specifies the relationship

between the alternatives and their attributes, and includes a vector of characteristics of individuals which condition the contribution of attributes to their utility. Choice probabilities are subject to two conditions; "choice probabilities are non-negative and sum to one, and choice probabilities depend only on measured attributes of alternatives and individual characteristics" (McFadden 1981, p. 201-202). The individuals make choices between alternatives, choosing the alternative which improves utility (a function of the attributes of the alternatives and individual characteristics).

In a binary choice model, where the individual is modeled as choosing Y_i, if the distribution is not continuous;

$$Y_i = 1 \quad \text{or}$$

$$Y_i = 0$$

The probability function describes the probability that an individual will choose alternative Y_i is given by the equation:

$$E(Y_i) = P(Y_i=1) = P_i = F(\beta' x_i)$$

where $0 \leq P_i \leq 1$ and $F(\beta' x_i)$ is a function of a set of explanatory variables. Conversely, the probability that $Y_i = 0$ is $1 - P_i$.

A specific example may help to clarify the equation described above. Consider an individual about to purchase a house where the choice of houses is limited. The utility of each alternative house depends on the price, size, number of bedrooms, location of the house as well as the income, number of children, and other characteristics of the decisionmaker. The probability that the individual selects a particular house is determined by the function $F(\beta' x_i)$. This function relates the characteristics of the alternatives and of the decisionmaker to the probabilities.

The alternatives in a probabilistic choice system make up the *choice set*. Each alternative included in the choice set for a specific individual must be feasible. That is, an alternative must be physically available, be known to the consumer and the consumer must possess the monetary resources necessary to select a particular alternative.

Individuals, when presented with repeated choices, frequently exhibit behaviors which are inconsistent and non-transitive. That is, individuals do not always select the same option even when faced with the same choice set. Also, observers have found that, if the choice set is altered, the assumption of transitive preferences[2] is often violated (Ben-Akiva and Lerman 1985). This observation leads to the presumption that a probabilistic choice system may better represent consumer preferences than does a deterministic system.

Consumers who possess identical socioeconomic characteristics, when faced with identical choice sets and choice attributes, sometimes make different selections. The different choices are attributed to a number of factors. These include the inability of the observer to see and measure all of the characteristics which affect behavior, the inherent probabilistic nature of decision making or the fact that the consumer may not perceive all of the choices correctly.

3.3 An Overview of Random Utility Theory

The theoretical basis for the model developed in this research is that of utility maximization where the consumer is expected to maximize satisfaction within the restraint of limited income. When this theory is applied to models of housing choice,

---

[2] Transitive preferences are present if when a is preferred to b, and b is preferred to c, then a is preferred to c.

the neoclassical model of utility maximization takes the form of:

$$maximize\ U = U(H,X);$$

$$subject\ to\ Y = P_h H + P_x X$$

where Y is income, P is price, H is quantity of housing services purchased and X is all non-housing goods purchased by the consumer. In that model, the choices available to the individual are housing, H, and the composite good, X. The variables, X and H, are assumed to be continuous.

In the normal consumer choice framework, the selection of housing is complicated by the fact that a consumer must select a single unit from a large variety of heterogeneous units with the differences attributable to housing characteristics, location and price (Quigley, 1985). In other words, housing choices involve the joint selection of a set of housing characteristics, a location and a price. In conventional marginalist theory, housing price carries all of the information about the good and a consumer has the ability to compare all possible alternatives. Choices are treated as continuous variables. Many of the dimensions of the housing choice are discrete (i.e. the tenure choice, the number of bedrooms). It is much more plausible to treat housing choice as a discrete variable and use the random utility function as a basis for modeling housing selection (Börsch-Supan and Pitkin 1988, McFadden 1981).

Random utility maximization (RUM) assumes that an individual's preferences are related to both the attributes of the alternatives and the characteristics of the individual. If all of the attributes and characteristics were known, it would be possible to predict choices deterministically (Ben-Akiva and Lerman 1985). The fact that they are not known with certainty suggests a probabilistic model. RUM is consistent with a probabilistic choice system and applicable to discrete choice analysis.

In random utility maximization, the utility function is divided into a deterministic (observed) component, V, and a random (unobserved) component, e. The form of the random utility function is: $U_{in} = V_{in} + e_{in}$. This is interpreted as the utility of selection i to individual n is the sum of the deterministic component and the random component. Including the random component makes the entire utility function probabilistic. The individual maximizes utility, but the utilities are not known with certainty to the analyst. RUM translates directly into a discrete choice framework.

3.4 The Discrete Choice Model

A discrete choice model of housing selection by low-income renters is developed in this section. The model is used to evaluate the probability that an individual household will select a particular combination of housing characteristics. In order to make the random utility model useful for computational purposes assume that;

1. Utility is divided into a deterministic and a random component.
2. Each component is specified.
3. The deterministic component, V, is that part of utility which is observed and the random component, e, is not directly observable.
4. Individuals are assumed to select the alternative that yields the highest utility, U. That is, given the two choice alternatives, i and j, U_{in} is selected if and only if;

   $$U_{in} > U_{jn} \quad i = j$$

   where i and j are two alternatives in a choice set. U_{in} is the utility derived by household n from the ith alternative and U_{jn} is the utility

derived by household n from the jth alternative.

Another way of looking at this is:

$$(V_{in} + e_{in}) > (V_{jn} + e_{jn})$$

on rearranging:

$$(V_{in} - V_{jn}) > (e_{in} - e_{jn})$$

The probability that household n will select alternative i equals the difference between the random utility of alternative i and alternative j. In discrete choice modeling, the variables which act as constraints on utility maximization in neoclassical choice theory are included in the utility function (Hensher and Johnson, 1981). The approach used is called the revealed preference approach because the utility function is inferred from observed choice behavior.

3.4.1 Specification of the Components

The deterministic and random components in a probabilistic choice system are the sources of utility. These components need to be specified prior to relating the selection probabilities to utility maximization.

The Deterministic Component. The alternatives in a choice set are characterized by their attributes. If a choice set of housing alternatives is considered, the attributes might include style, number of rooms or size and lot size. Since there are a number of attributes, they are contained in vectors[3] of attributes, z_{in} and z_{jn}. Similarly, the characteristics of the decision maker (here a household) are numerous and contained in another vector, S_n. For the sake of clarity, both the vector of attributes and the vector of socioeconomic characteristics are combined in a single vector, X_n.

[3] This notation is directly from Chapter 3 of Ben-Akiva and Lerman (1985).

In the binary choice situation, the deterministic components of the random utility function are represented by V_{in} and V_{jn}. The problem of specification becomes one of specifying the relationship between z_{in} and S_n to represent V_{in}, and between z_{jn} and S_n to represent V_{jn}. This relationship is defined by a new vectors, x_{in} and x_{jn}. The deterministic components of utility are now written as:

$$V_{in} = V(x_{in}) \text{ and } V_{jn} = V(x_{jn}).$$

The functional form of the deterministic component, V, is generally assumed to be linear in the parameters (Ben-Akiva and Lerman 1985, p. 63). This permits the deterministic component to be specified as:

$$V_{in} = \beta_{i0} + \beta_{i1} + \ldots\beta_{ik}$$

and

$$V_{jn} = \beta_{j0} + \beta_{j1} + \ldots\beta_{jk}$$

where β_{in} and β_{jn} are vectors of parameters.

Ben-Akiva and Lerman (1985, p. 3-64) make two important points regarding the parameters. First, linearity in the parameters does not imply that the attributes of the alternatives or the characteristics of the decision maker have to be linear. Second, there is an implicit assumption that the parameters $\beta_1,\beta_2,\ldots\beta_k$ are identical for the entire population. This implies that separate sub-models are needed for various population groups only if there is reason to suspect that different groups have different preferences or tastes. To the extent that tastes "vary systematically with the observed variables" (Train 1986, p. 29) the logit model captures taste variations. The logit model cannot capture random taste variations or taste variations which vary with unobserved variables.

Anas (1982) suggests that the population of households consists of a number of segments where socioeconomic characteristics identify each segment. Estimating

choice models for each population segment avoids bias (p. 123). Income level and urban location identifies the population segment used in this research. The models include only those urban households falling at or below eighty percent of median income. This segmentation reflects the focus of the dissertation on low-income households who are generally eligible for housing assistance.

The Random Component. The aim of a discrete choice model is to estimate the significance of the determinants of V. However, the distribution of the random components makes the model probabilistic and assumptions about the shape of this distribution influence the selection of a statistical model.

Recall the equation above which divided utility into deterministic and random components. The rearranged equation was $(V_{in} - V_{jn}) > (e_{in} - e_{jn})$. Here, the difference is relevant not the absolute value. Assumptions about the shape of the distribution of the e_{in} and e_{jn} terms or about the distribution of $(e_{in} - e_{jn})$ leads to the selection of a particular statistical model. The only consistent assumption is that the random components are independently and identically distributed (IID).

Since the random components are not observed by the analyst, there is no *a priori* method for determining the shape of the distribution. Housing choice researchers have generally assumed that these disturbances follow either a normal distribution, a logistic distribution or a generalized extreme value distribution[4]. Assuming a normal distribution leads to the estimation of a probit model. The logistic distribution leads to a logit model and the generalized extreme value distribution leads to variations of the multinomial logit model.

The analysis in this research is directed at a series of binary choices. Therefore, this discussion addresses only binary choice models. By limiting the

[4] For an excellent discussion of these distributions see McFadden (1981).

analysis to binary choices, the assumption of the independence from irrelevant alternatives (IIA) does not need to be considered. Briefly, the IIA assumption states that "the ratio of the probabilities of choosing one alternative over another is unaffected by the presence or absence of any additional alternatives in the choice set (Hensher and Johnson, 1981 p. 38)." In other words, relative probabilities between two alternatives do not change when other alternatives are added or removed from the choice set. Dispensing with the IIA assumption is possible in this research because each choice set contains only two alternatives. This greatly simplifies the modeling process and the subsequent analysis.

3.4.2 The Logit Model

The probit model and the logit model are both used in binary choice situations. The probit model is the logical choice if the researcher assumes that the random components, e_{in} and e_{jn}, are distributed normally about a mean of zero with variances, σ_i^2 and σ_j^2. The probit model expresses choice probabilities as integrals and interpretation is difficult. The logit model is a more attractive choice for most researchers. The assumption that the random components are distributed logistically leads to its derivation. The error terms, e_{in} and e_{jn}, are independent and identically Gumbel[5] distributed with a variance of $\pi^2/6$ (Ben-Akiva and Lerman 1985, p. 71). This distribution approximates a normal distribution except that the tails are fatter.

Binary Logit. Recall that:

$$E(Y_i) = P(Y_i=1) = P_i = F(\beta' x_i)$$

[5] The properties of the Gumbel distribution can be found in Ben-Akiva and Lerman (1985 p. 104-106). The necessary assumption for the binary logit model is that $e_n = e_{in} - e_{jn}$ is logistically distributed.

This model relies on the assumption that the F is the cumulative logistic probability function[6].

The form is:

$$P_i = e^{\beta' x_i} / [1 + e^{\beta' x_i}]$$

and

$$(1 - P_i) = 1 / [1 + e^{\beta' x_i}]$$

The probability ranges from 0 to 1 as $(\beta' x_i)$ ranges from $-\infty$ to $+\infty$ and the linear expression becomes:

$$P_i [1 + e^{\beta' x_i}] = e^{\beta' x_i}$$

$$P_i = (1 - P_i) e^{\beta' x_i}$$

$$P_i / (1 - P_i) = e^{\beta' x_i}$$

$$\log_e [P_i / (1 - P_i)] = (\beta' x_i)$$

The left-hand side of this equation is called the logit or odds ratio.

3.4.3 Maximum Likelihood Estimation of the Logit Model

Discrete choice models do not fit into the category of models which are estimated using ordinary least squares techniques. Because the dependent variable is categorical not continuous, alternative methods for estimating the parameters are needed. The method most commonly used is maximum likelihood (ML). This is an optimization procedure which relies on methods of non-linear optimization to obtain the parameters of the function. A commonly used method is the Newton-Raphson method[7].

The method of maximum likelihood is based on the chance or likelihood of

---

[6] This derivation is from Wrigley (1985) p. 28-29.

[7] Readers familiar with this type of modeling may prefer to skip to Section 3.4.4.

observing patterns of responses occurring in a data set. A process is then used to estimate the unknown parameters $(\beta_0,\beta_1...\beta_k)$ which maximize the likelihood of the pattern of responses. Wrigley (1985, p. 35) illustrates the development of maximum likelihood estimators using a binomial distribution with the likelihood of obtaining a specific sample of responses determined by the joint probability of occurrence of the actual responses.

The likelihood function is:

$$L(\beta_0,\beta_1,...\beta_k) = \prod P_i$$
$$= \prod e^{\beta' x_i} / (1 + e^{\beta' x_i}) \prod 1 / (1 + e^{\beta' x_i})$$

To solve the above equations in an efficient manner, the partial derivatives must be determined. A more direct method is to transform the likelihood function into its logarithmic form. The resulting log form is easier to differentiate. The function then becomes:

$$\log_e = L(\beta_0,\beta_1,...\beta_k)$$
$$= \Sigma \log_e (\beta' x_i) - \Sigma \log_e (1 + e^{\beta' x_i})$$

Hence, the maximization problem is to maximize the log of the likelihood function. McFadden (1974) shows the complete derivation of the function and proves that the function is concave with a strict maximum. The maximum likelihood estimators are sufficient and distributed asymptotically normal. Asymptotic normality refers to properties estimators produced by this method. These properties, unbiasedness, consistency, normality and efficiency, apply only to large samples.

3.4.4 Measuring Goodness-of-Fit and the Influence of Individual Variables

This paper uses a number of tests to assess model fit. As an overall

evaluation measure, most recommend the pseudo R^2 or rho-square, p^2 (McFadden 1974, Ben-Akiva and Lerman 1985, Hensher and Johnson, 1981). This measure is determined as follows:

$$p^2 = 1 - [(\log L(\beta_k) / (\log L(\beta_0)]$$

or one minus the ratio of the log likelihood values at the maximized value to the log likelihood with the constant only term included. It is generally believed that this ratio exhibits a good fit if it lies between 0.2 and 0.4 (McFadden 1974, p. 24). An adjusted version of this measure is also used, where the adjustment is for the number of parameters in the model.

Some researchers use a prediction success table as a overall measure of the model. This measure matches the observed responses with the predicted responses. Ben-Akiva and Lerman (1985, p. 92) caution that overreliance on this measure is not desirable because it can "mask poor goodness of fit." They favor the p^2 as the preferred measure. Another problem associated with using percent correctly predicted is that prediction success tables use a statistic which relates a sampled decisionmaker with the alternative with the highest probability. Train (1986) argues that this statistic is not consistent with the reason for specifying choice probabilities. The researcher does not have enough information to predict a decisionmakers choice every time but only state that if the situation were repeated many times there is a probability associated with making a specific choice.

In evaluating the model, it is also important to examine the individual parameters. A frequently used measure to test the null hypothesis that all of the parameters except the alternative specific constant are zero is:

$$-2 [\log L(\beta_0) - \log L(\beta_k)]$$

where $\log L(\beta_0)$ is the value of the log likelihood when all the parameters are zero

and $log\ L(\beta_k)$ is the value of the maximized log likelihood function. This statistic is distributed chi-square with k - 1 degrees of freedom.

Individual parameters are evaluated using an asymptotic t statistic. This is calculated by dividing the parameter estimate by its corresponding standard error. The statistic is distributed asymptotically normal. This means that it is only reliable for large samples.

3.4.5 Discussion

The measures discussed in Section 3.4.4 present some meaningful ways of measuring model fit. They do not assess the predictive power of a model. Frequently, in order to validate the model, data is partitioned into two subsets. One subset is used as an estimation data set and the second is used as a validation data set. In this way, a meaningful comparison of the estimated vales to the fitted values is made. This technique, however, is difficult to use in the type of hierarchical model presented in this paper. The requirement that the data be subsetted at each sucessive level of the hierarchy results in reduced sample sizes. If the data were partitioned into two data sets initially, the sample would be reduced below the level necessary for estimating the model. Ben-Akiva and Lerman (1985, p. 372) point out that many advances in modeling processes yield consistent and asymptotically efficient estimators. They add, however, that this type of modeling is still rooted in behavioral theory which has changed very little in the recent past. Housing selection behavior is complex and diffficult to model. Until more resources are directed at exploring behavioral aspects (i.e. satificing rather than maximizing behavior) of housing selection, the utility maximization hypothesis produces the most tractable results.

3.5 Representing Sources of Utility in the Choice Model

Hensher and Johnson (1981) see the lack of clearly defined empirical criteria for determining the elements in a choice set as a problem in model estimation and application. They point out that an individual may have a different interpretation of the alternatives than the definitions of alternatives used by the researcher. When the number of alternatives is large or data for all alternatives is not available, researchers tend to group alternatives into subsets or select representative alternatives to reflect the full range. The groups or representative alternatives may not reflect the decision makers view of a choice set.

To some degree, the use of a binary choice set which simply records the presence or absence of an individual in a categorical group masks problems associated with grouping elemental alternatives. It does not, however, mask definitional problems. For example, crowding is defined by a persons per room indicator with more than one person per room being defined as crowded. A particular individual (household) may not perceive a housing unit as crowded until the number of persons per room exceeds two or even more. Or a young couple may perceive their three room apartment as crowded when a baby is born even though the persons per room indicator suggests that it is not.

The model presented here assumes that utility arises from alternatives in the choice sets estimated in this research. That is, utility arises from the attributes of affordability, quality and crowding which were discussed in Chapter 2. The choice set needs to be segmented into feasible and nonfeasible choice subsets for the purpose of empirical investigation. By simplifying the choice set into a series of binary decisions, I assume that every household in the sample faces a market where

these choices exist. Ideally, market data would be used to verify this assumption. The AHS data does not include any market specific information. However, even in the absence of such information, the assumption seems reasonable due to the diverse nature of housing in urban areas.

3.5.1 Affordability

The first model estimated is the affordability sub-model. This model is used to test hypotheses about using the thirty percent rent-to-income ratio as a measure of the housing cost burden faced by low-income households. Recent evidence (Stone 1990) suggests that such ratios understate affordability problems for large households and overstate affordability problems for higher income households. Clearly, the arbitrary nature of a categorical variable does not reveal substantial information about the underlying continuous variable. Housing programs, however, traditionally rely on such measures to determine assistance needs. This variable takes on a value of one if the rent burden exceeds thirty percent of income and zero otherwise.

A variety of factors affect the relative cost of housing facing each household. Thus for an affordability choice, the specification of the equation might include socioeconomic and demographic variables and a variable to proxy market conditions. Socioeconomic and demographic variables are easily determined from the AHS data. The inclusion of a variable called the alternative specific constant compensates for the lack of market information. The purpose of an alternative specific constant (ASC) is to capture effects of unobserved factors which affect choice (Hensher and Johnson 1981)[8]. The ASC variable takes on a value of one for

[8] The alternative specific constant is often compared to the intercept term in an ordinary least squares regression equation. In a binary logit, it is the mean of the difference in the error terms.

a given alternative in a binary model and zero otherwise. The resulting choice set reflects a combination of underlying choice dimensions.

3.5.2 Quality

An additional decision which a household makes in the selection of housing is that of quality. Obviously, housing quality is not easily captured by a binary variable. Yet, inadequate housing is associated with having a specific number and type of housing problems. Degrees of inadequacy are not considered by housing assistance programs. Hypotheses about the extent to which a binary classification of adequacy reflects actual choices made by low-income households are tested by the quality model.

The quality decision is represented by a dichotomous variable which takes on a value of one if the unit is judged to be of inadequate quality and zero otherwise. The measurement of quality follows that used by Newman and Schnare (1989) and is completely described in Appendix A. At this decision level, the choice set is partitioned based on the observed affordability decision. In addition to socioeconomic and demographic variables, the model includes an alternative specific constant and variables related to judgments about quality are included in this model.

3.5.3 Crowding

The crowding function is estimated in a manner similar to the quality variable. The variable takes on a value of one if the number of persons per room exceeds one and zero otherwise. The crowding function also includes an alternative specific constant. A final partitioning of the data set occurs prior to estimating this

model. The basis for this last partitioning is the observed combination of affordability and quality.

3.6 The Hierarchy of Housing Choices

The three equations described above are combined to form a housing choice hierarchy. By doing this, I assume that a hierarchical relationship exists between this set of interdependent alternatives or that the housing decision entails the sequential choice elements which comprise a complex alternative. Boehm (1979) points out that the order of a hierarchy is difficult to defend theoretically and suggests estimating the model with different orderings of the choices. In this research, I assume that the choice hierarchy of affordability first, quality second and crowding third holds. The justification for this assumption lies in the distribution of housing problems among the low-income population. The most prevalent problem for low-income households is a high rent burden. This is followed by quality problems and finally crowding problems (see Table 4.1).

The first level of the hierarchy is hypothesized to contain the estimation of a household's affordability decision, $P(A)$. The quality decision is estimated at the second level, $P(Q \mid A)$. This decision is contingent upon the affordability choice. At the third level, the crowding decision is made, $P(C \mid Q,A)$. This decision is contingent on the previous two choices. Following the modeling technique used by Boehm (1979), the choice equations are combined to form joint probabilities. It is possible to view this hierarchical choice model in the form of a decision tree, Figure 3.1. Here, the affordability choice is on the first level and subsequent choices follow.

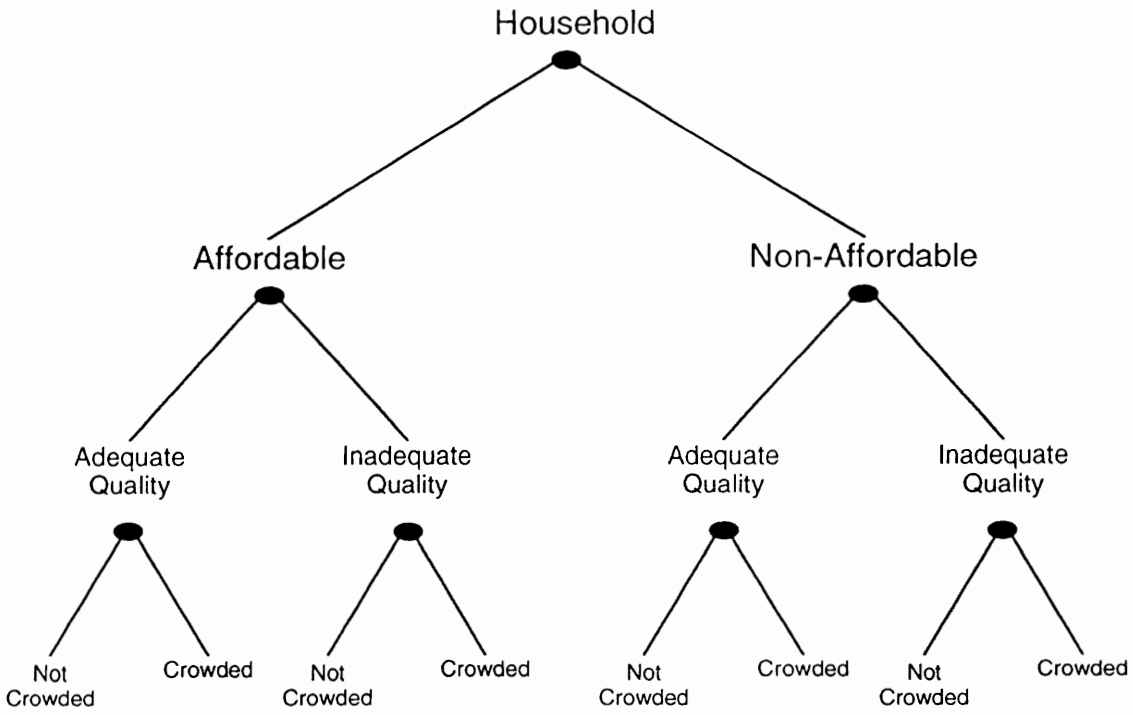


Figure 3.1 Housing Choice Hierarchy

Recall that the affordability function, A, has two levels representing rent-to-income ratios at or below thirty percent, A, and above thirty percent, NA. The quality variable Q has two levels AQ, adequate quality, and IQ, inadequate quality. Finally, the crowding variable, C, is defined as NC, not crowded, and OC, crowded. The resulting probabilities are[9];

$$P(A,AQ,NC) = P(A) \bullet P(AQ \mid A) \bullet P(NC \mid AQ,A)$$

$$P(NA,AQ,NC) = (1-P(A)) \bullet P(AQ \mid NA) \bullet P(NC \mid AQ,NA)$$

$$P(A,IQ,NC) = P(A) \bullet (1-P(AQ \mid A)) \bullet P(NC \mid IQ,A))$$

$$P(NA,IQ,NC) = (1-P(A)) \bullet (1-P(AQ \mid NA)) \bullet P(NC \mid IQ,NA))$$

$$P(A,AQ,OC) = P(A) \bullet P(AQ \mid NA) \bullet (1-P(NC \mid AQ,A)$$

$$P(NA,AQ,OC) = (1-P(A)) \bullet P(AQ \mid NA) \bullet (1-P(NC \mid AQ,NA))$$

$$P(A,IQ,OC) = P(A) \bullet (1-P(AQ \mid A)) \bullet (1-P(NC \mid IQ,A))$$

$$P(NA,IQ,OC) = (1-P(A)) \bullet (1-P(AQ \mid NA)) \bullet (1-P(NC \mid IQ,NA))$$

I estimated the parameters of the first level, affordability, from the entire sample of households. Following that, I divided the sample into two groups, those households who live in affordable housing and those households who do not. Estimating the parameters of the second level, quality, followed and the sample was divided again. This final division was into four groups, one for each combination of affordability and quality. Estimating crowding at the third level completed the process. A model of this type is tractable and not difficult to use under the assumption that the random factors which influence decisions at each level are independent (Maddala 1988, p. 50). The structure assumes that each decision is

---

[9] The notation: $P(A \text{ and } AQ \text{ and } NC) = P(A) \bullet P(AQ \mid A) \bullet P(NC \mid AQ,A)$ is read as: the probability of affordable, adequate quality, uncrowded housing is equal to the probability of affordable housing times the probability of adequate quality housing given that the unit is affordable, times the probability of uncrowded housing given that the unit is both adequate quality and affordable.

conditional on the previous decision. There are no feedback loops or mutual interactions (Hensher and Johnson 1981).

The picture of housing deprivation, then, is a composite of housing affordability problems, housing quality problems and the presence or absence of crowding. A critical factor in understanding how discrete choice modeling relates housing problems to housing policy is understanding the process by which households are categorized by housing programs. Individual thinking about housing problems is likely to produce vastly different conclusions. By choosing these particular measures, housing policy may not reflect the nature or extent of the housing problems of low-income households. Following a description of the data presented in Chapter 4, Chapter 5 develops an empirical model to test hypotheses about these measures.

CHAPTER 4
THE DATA

This chapter describes the data used to estimate the model developed in Chapter 3. The data are from the national sample of the 1989 American Housing Survey (AHS)[1]. The 1989 National Core file contains data on 58,942 housing units. The housing unit, not the occupants, is the focus of the survey. However, extensive information is compiled on both. The national sample is based on randomly selected units from the 1980 Census. Interviewers from the Census Bureau visit each unit in the sample. If the unit is occupied, the interviewer obtains information on both the unit and the household occupying it. If the unit is vacant, the interviewer obtains information on the unit from the landlord, rental agent or neighbor. This research includes only occupied units.

The next section presents a description of the sample of low-income renters used to estimate the discrete choice model. This is followed by selected statistics on the composition of the low-income renter population. The final section relates the composition of low-income households to the presence or absence of assistance.

---

[1] For a complete description of the AHS data, see the AHS Codebook (Hadden and Leger 1989).

4.1 The Recent Mover Sample

In order to estimate the hierarchical model described in Chapter 3, I used a sample of low-income households who moved recently. Recent movers are usually assumed to be at or near an equilibrium level of housing consumption[2]. A large initial sample is required because the sample is reduced as each level of the hierarchy is estimated. Two more limiting factors are the requirement that all households be renters and that income falls at or below eighty percent of median income for the region. As pointed out in the discussion of errors in AHS data in the AHS Codebook (Hadden and Leger 1989), income is frequently underreported. Since there is no way to verify underreporting, I assumed that households spending more than one hundred percent of their income on housing are spending out of savings and retained them in the study.

Bohem (1979) points out a further consideration in estimating housing choice models. The model assumes implicitly that the household is fully informed regarding the housing choices available when they move. Two types of recent movers are found in the AHS data, those who make intra-metropolitan moves and those who make inter-metropolitan moves. It is reasonable to assume that both types of movers lack full market information. Inter-metropolitian movers may make more rational choices unconstrained by random factors such as past location, school preferences or the location of friends. However, including only one type of mover reduces the sample below a meaningful level. Therefore, the final sample includes all recent, low-income movers living in an urban area with an assumption of typical market information.

---

[2] Recognizing that low-income households' reasons for moving differ from households with higher incomes, I use the word equilibrium somewhat cautiously. Arguably, many low-income households are faced with constant disequilibrium moving to avoid disaster not reach equilibrium.

4.1.1 The Composition of the Recent Mover Sample

Fortunately, the AHS data set is of sufficient size so that the limiting factors mentioned above do not affect the initial levels analysis. As shown in Chapter 5, however, the sample becomes quite small when the third level of the hierarchy is estimated. There are 16,663 renter households in the 1989 AHS data set. Of those households 10,399 are considered low-income renters. Low-income households are those whose income falls at or below eighty percent of their regional median income. Two-hundred and eighty observations were eliminated from this group for miscellaneous reasons[3]. Of the remaining households, 6,644 did not move during the previous twelve month period leaving a sample of 3429 mover households. The sample used in the logit models in Chapter 5 includes only 3238 households. Missing data caused the elimination of the additional 191 households.

4.1.2 Description of Data

Table 4.1 illustrates the housing affordability measure (rent-to-income ratio) for all renters, low-income renters and the recent mover sample. As other studies have shown, rent-to-income ratios are normally higher for low-income renters (Newman and Schnare 1988, Stone 1990, Feins and Saunders 1981) than for the entire renter population. The mean rent-to-income ratio for renters is in Table 4.2. The affordability burden is most notable in the recent mover sample. A possible explanation is that recent movers rent reflects current market conditions (no time related discounts). The higher ratio does not occur because the income of recent movers is lower.

---

[3] These were rural households, households living in public housing projects or households with privately supplied housing. I eliminated these groups in order to have a sample of households who must face finding housing in an urban area at market rental rates.

Table 4.1

Housing Affordability by Region and Income Level
Rent-to-Income Ratio 1989 (Percent)

All Renters	North East	North Central	South	West	
R/I < = 30%	60.04	63.41	67.71	59.38	
30% - 50%	21.60	20.05	21.62	24.18	
R/I > 50%	18.36	16.55	14.67	16.47	
n =	3673	3342	4478	3648	Total 15144
Low-Income Renters					
R/I < = 30%	43.46	50.00	45.28	39.97	
30% - 50%	29.87	27.13	31.61	34.26	
R/I > 50%	26.67	22.87	23.11	25.77	
n =	2531	2418	2838	2332	Total 10119
Low-Income Recent Movers					
R/I < = 30%	34.49	45.65	40.80	35.34	
30% - 50%	32.95	29.80	31.46	37.20	
R/I > 50%	32.56	24.55	27.74	27.46	
n =	519	839	1157	914	Total 3429

Source: 1989 American Housing Survey

Table 4.2

Mean Rent-to-Income Ratio by Region 1989 (Percent)

	North East	North Central	South	West
All Renters	33.78	32.19	31.10	32.97
Low-Income	41.76	38.90	39.27	41.31
Recent Movers	46.10	40.87	42.32	42.93

(No public ownership)
Source: 1989 American Housing Survey

Table 4.3 presents the distribution of both groups of low-income renters by region and by income with respect to medium income. The most noteworthy aspect of this table is that the income distribution of all low-income renters does not differ substantially from the recent mover group. The greatest difference between the two groups is about five percent. A larger proportion of recent movers are in the fifty to eighty percent of median income group than are low-income renters in general.

Since this study uses data on recent movers, we need to ask how representative is that group relative to all low-income renters. Do recent movers reflect the total low-income population? Table 4.4 presents the distribution of all low-income households by housing problem and region and separately shows the same information for the recent movers sample. In every case, the proportion of recent movers spending more than thirty percent of their income on housing exceeds the proportion of the total low-income group. However, quality problems are not consistently different between groups. Where differences exist, the amount is small. It does appear that recent movers are somewhat more crowded than are all low-income renters. A greater rent burden coupled with more crowding suggests that recent movers are worse off than the total low-income group.

Housing assistance is a way to mitigate housing problems. It is supposed to reduce affordability problems and insure that some type of minimum quality and crowding standards are met. As seen in Table 4.5, the percentage of low-income households receiving any type of assistance is extremely small. The northeast section of the country receives the highest level of housing assistance, both for the total low-income sample and the recent mover sample.

Table 4.3

Low-Income Renters by Region and
Percent of Regional Median Income

	North East	North Central	South	West
50% -80%				
All Low-Income	35.91	31.02	34.14	40.09
Recent Movers	40.08	35.16	37.86	43.98
Less than 50%				
All Low-Income	64.09	68.98	65.86	59.91
Recent Movers	59.92	64.84	62.14	56.02

Source: 1989 American Housing Survey

Table 4.4

Low-Income Renters by Location and Housing Problem

All Low-Income[a]	North East	North Central	South	West
Affordability				
R/I < = 30%	43.46	50.00	45.28	39.97
R/I > 30%	56.54	50.00	54.72	60.03
Quality				
Adequate	76.29	78.70	68.32	60.21
Inadequate	23.71	21.30	31.68	39.79
Crowding				
PPR < = 1	96.01	96.20	94.89	90.78
PPR > 1	3.99	3.80	5.11	9.22
Recent Movers [b]				
Affordability				
R/I < = 30%	34.49	45.65	40.80	35.34
R/I > 30%	65.51	54.35	59.20	64.66
Quality				
Adequate	76.49	76.64	70.10	62.91
Inadequate	23.51	23.36	29.90	37.09
Crowding				
PPR < = 1	95.57	94.99	94.64	89.61
PPR > 1	4 .43	5.01	5.36	10.39

[a] n = 10119
[b] n = 3429

Source: 1989 American Housing Survey

Table 4.5

Percent of Low-Income Renters Receiving Housing Assistance by Region

All Low-Income[a]	North East	North Central	South	West
Federal	5.73	5.17	6.45	4.97
Local	2.17	2.11	0.88	2.32
Recent Movers[b]				
Federal	6.17	3.22	5.01	4.60
Local	1.93	2.38	0.78	2.30

[a] n = 10119
[b] n = 3429

Source: 1989 American Housing Survey

4.2 Summary

American Housing Survey data indicate that in 1989 the average low-income household in the South experienced more housing problems and received less housing support than did low-income households in other parts of the country. This pattern will probably continue unless the distribution of housing assistance undergoes substantial change. The uneven allocation of assistance across regions reflects both local decisions and federal budget decisions. The empirical analysis in the next chapter further addresses differences between those who receive assistance and those who do not.

CHAPTER 5
EMPIRICAL RESULTS

This chapter presents the results of estimating the model described in Chapter 3. The model, a hierarchical model of housing choice, uses a sequence of three different binary logit models to analyze the revealed housing decisions of low-income urban renters. Section 5.1 describes the construction of the dependent variables used in each of the binary logit models. Section 5.2 considers the independent variables. A concise description of the independent variables is presented in Table 5.1. The results of the empirical analysis are presented in Section 5.3. Section 5.4 discusses the joint selection of affordability, quality and crowding by three representative household types. The final section, Section 5.5 offers some concluding comments.

5.1 Dependent Variable Description

As noted in the introduction, a primary goal of this paper is to illustrate the use of commonly held indicators of housing problems as elements in the choice set facing low-income urban renters. The usual indicators of housing problems are related to affordability (A), quality (Q) and crowding (C). All of the logit models

assume binary dependent variables where the associated probability is that a household will live in affordable *P(A)*, adequate quality *P(AQ | A)*, uncrowded *P(NC | A,AQ)* housing. Therefore, in discussing the expected signs, the expectation is that a positive sign increases the probability and a negative sign reduces the probability.

Affordability. Affordability is a binary variable constructed by taking the total annual costs of renting and dividing that number by the household's total annual income. The total annual cost of renting (full rent) includes both contract rent and utilities. A housing unit is classified as affordable if the resulting rent-to-income ratio is at or below thirty percent. That is, a household spending up to thirty percent of total income on housing is presumed to meet a test of affordability. If the household spends more than that amount of income on housing, the housing is classified non-affordable. This indicator does not consider either the income level of the household nor the related housing market. A number of housing analysts[1] point out the arbitrary nature of this indicator, however, it continues to be used in determining housing policy.

Adequate Quality. This variable, adequate quality, reflects the accumulation of particular housing problems. It is a binary variable categorizing a housing unit as either adequate or inadequate quality. Construction of the variable required evaluating a number of items in the American Housing Survey data set. The design of this variable, based on a modified HUD index, follows that used by Newman and Schnare (1988)[2]. Newman and Schnare describe the index as ranking high in external validity. That is, when this index ranks a housing unit as inadequate

---

[1] See for example Feins and Saunders 1981 or Stone 1990.

[2] For a complete description of the construction of this variable see Appendix A.

quality, then the housing unit appears as inadequate on other indexes of quality (Newman and Schnare 1988, p. 99). In comparing this index to nine other commonly used indexes, they found "roughly comparable" results from six indexes.

Not Crowded. The crowding variable relates the number of persons in a household to the number of rooms in the dwelling. The variable is the number of persons in the household divided by the number of rooms. If the resulting ratio, persons per room (PPR), is one or less then the unit is not crowded. If the ratio exceeds one the unit is reported as crowded. Crowding is not very prevalent in the AHS data even among the lowest income levels.

5.2 Independent Variables

The housing choices made by low-income urban renters are functions of social and economic characteristics. There is general agreement that these two types of variables fall into broad key categories. The categories used in this analysis are:

- Demographic
- Socioeconomic
- Housing
- Geographical

A discussion of each of the categories and the specific variables used in the analysis follows. Table 5.1 serves as a quick reference to the independent variables. Table 5.2 presents simple statistics for each of these variables for the sample of households used in the model.

TABLE 5.1
Independent Variable Description

Demographic Variables	
RACE	Race of household head, 1 = white
SEX	Sex of household head, 1 = female
PERSONS	Number of persons in the household
MARRIED	Marital status of household head, 1 = married
AGE	Age of household head
GRADE	Education of household head. Eighth grade or less is GRADE < =8, between eighth grade and twelfth grade is GRADE < =12 and more than high school is the third group.
Socioeconomic Variables	
INCOME	Total household income in thousands
FEDERALLY ASSISTED	Federal housing assistance, 1 = yes
LOCALLY ASSISTED	Local housing assistance, 1 = yes
Geographic Variables	
NORTH EAST NORTH CENTRAL SOUTH WEST	Regions of the country.
CENTRAL CITY	Relative location within the urban area, 1 = central city.
Housing Related Variables	
MULTI-FAMILY	Housing unit is classed as multi-family, 1 = multi-family
UNIT AGE	Approximate age of dwelling unit. UNIT AGE < 10 are units less than ten years old, UNIT AGE < 20 are from ten to twenty years and UNIT AGE > 20 are greater than twenty years.

Table 5.2
Simple Statistics for Independent Variables

Demographic Variables	
RACE	22% Non-White 78% White
SEX	53.10% Male 46.90% Female
PERSONS	Mean 2.3684 Std. Dev. 1.5274
MARRIED	71.70% Not Married 28.30% Married
AGE	Mean 34.0299 Std. Dev. 14.1313
GRADE	8.80 % GRADE <=8 54.50 % GRADE <=12 36.60 % GRADE > 12
Socioeconomic Variables	
INCOME	Mean $14,336 Std. Dev. $7,528
FEDERALLY ASSISTED	95.27 % No 4.73 % Yes
LOCALLY ASSISTED	98.21 % No 1.79 % Yes
Geographic Variables	
NORTH EAST	15.23 %
NORTH CENTRAL	24.09 %
SOUTH	33.57 %
WEST	27.12 %
CENTRAL CITY	48.92 % No 51.08 % Yes
Housing Related Variables	
MULTI-FAMILY	27.70 % No 72.30 % Yes
UNIT AGE	21.20 % UNIT AGE < 10 44.70 %, UNIT AGE < 20 34.10 % UNIT AGE > 20

5.2.1 Demographic Variables

Many demographic factors are potentially relevant in evaluating housing choice models. The issue of taste frequently justifies the inclusion of particular demographic variables in housing models. That is, a large family may have a *taste* for more housing (Goodman 1990). The set of demographic variables found in discrete choice studies commonly includes age, sex, race, marital status and education of the household head. Also, the number of persons in the household and the presence of children under the age of eighteen are frequently included. Goodman (1990) identifies a number of studies using these variables and their impact on housing demand equations. While most housing choice research relates demographic variables to the tenure decision (Henderson and Ioannides 1987, Börsch-Supan and Pollakowski 1986), this study uses demographic variables to evaluate the probability of making decisions related only to rental housing.

5.2.2 Socioeconomic Variables

The socioeconomic variables included in this study relate to household income and to the presence of and type of housing assistance[3]. Housing opportunities are certainly greater for relatively wealthier households, however, less wealthy households are less constrained in finding housing if they receive some type of housing assistance. Since the sample is restricted to households with incomes at or below eighty percent of the regional median, these households are generally eligible for housing assistance. However, very few households actually receive any assistance. Households not receiving assistance should have a lower probability of occupying affordable, adequate quality, uncrowded housing than assisted

[3] As noted in the description of the sample (Chapter 4), no public housing or employer provided housing units are included.

households.

The income variable, INCOME, is annual household income in thousands. It includes the income from all sources of all household members including non-relatives. Scaling the income variable in thousands produced more reasonable parameter estimates. Very low-income renters often pay more than half of their income for rent (Nelson and Khadduri 1992). However, even among the low-income population problems decline as income increases. More than half of the population in the sample are very low-income with incomes at or below fifty percent of regional median.

Housing assistance either serves to increase income or to reduce the cost of housing. It acts as an income supplement when the form of assistance allows the use of more income for other forms of consumption. A housing voucher which lets the household spend less than thirty percent of income on housing[4] is an example of this type of assistance. The Section 8 program which covers the difference between the contract rent and thirty percent of income is the type of assistance which effectively reduces the cost of housing. A binary variable, FEDERALLY ASSISTED, indicates the presence of federal housing assistance. Unfortunately, the AHS data does not indicate the type of federal housing assistance that the household receives. LOCALLY ASSISTED indicates the receipt of local housing assistance. Local housing assistance varies greatly from community to community and region to region (Newman and Schnare 1988). Both variables are included to help evaluate the contribution housing assistance makes to a housing outcome. However, in order to only look at choices facing households in the marketplace, the sample does not

[4] Housing vouchers are for a specific dollar amount which is the difference between thirty percent of income and rent on a typical unit in the area. A household may spend less than thirty percent of income on rent if a unit for less can be located. Alternatively, a household may choose to spend more.

include publicly owned or privately supported (e.g. parsonages) units.

5.2.3 Housing Variables

The analysis uses four housing related variables. These variables reflect factors facing the household which the household presumably considers when making a housing choice decision. There is general agreement that, in the United States, households prefer single family housing to multi-family housing. A variable, MULTI-FAMILY, is include to aid in evaluating the role that preference plays in selecting housing. The variable takes on a value of zero if single family housing and a value of one if multi-family (including duplex through high-rise).

Categorical variables identify the approximate age of the dwelling. These variables occur only in the quality sub-model. UNIT AGE < 10 identifies units which are less than ten years old. UNIT AGE < 20 identifies units from ten to twenty years old and UNIT AGE > 20 dwellings older than twenty years. UNIT AGE > 20 occurs when UNIT AGE < 10 and UNIT AGE < 20 are both equal to zero. Housing quality generally declines as a unit ages. This variable captures the extent to which that decline influences quality choice.

5.2.4 Geographic Variables

Income levels and housing markets vary considerably in different parts of the United States. Although still a relatively gross measure of this variation, the data set identifies four multi-state regions. These are the categorical divisions used by the census as well as the AHS. They are designated as follows; NORTH EAST, NORTH CENTRAL, SOUTH, and WEST. As with typical dummy variables, WEST occurs when NORTH EAST = 0, NORTH CENTRAL = 0, and SOUTH = 0.

The analysis includes a second geographic variable, CENTRAL CITY, indicating the unit's location within the urban area. The location variable divides a metropolitan area into two categories, central city and suburb. The designation follows that of the AHS with all suburban categories being classed suburb. Residents with the lowest opinion of their neighborhood are generally central city residents living in multi-family dwellings (Boehm and Ihlanfelt, 1991). Households may live in center city areas because housing is more affordable. The location variable, coded one for central city, is a way to evaluate how living in a center city area affects affordability, quality and crowding.

5.3 Estimation Procedure

This research used a sample of 3238[5] low-income urban households who moved from one dwelling to another during the prior twelve month period to estimate the parameters of the model. Calculating the probability that a household selects a particular combination of affordability, quality and crowding required estimating three separate logit sub-models. The first model estimated was the affordability model. Estimation procedures involved first fitting a binary logit model to the entire data set and subsequently fitting six additional binary logit models conditioned on previous choices. Figure 5.1 is a graphic illustration of the housing choice hierarchy. Each node in the illustration represents a binary decision made by the household. The binary logit model is in the form of $log_e [P_i / (1 - P_i)] = (\beta' x_i)$ or the log odds that the revealed preference of respondent i is a function of a set of independent variables x. This is expressed as the ratio of the probabilities. These models reflect the probability that a household is able to find an affordable,

[5] Of the sample of recent movers discussed in Chapter 4, 191 households were eliminated due to missing data.

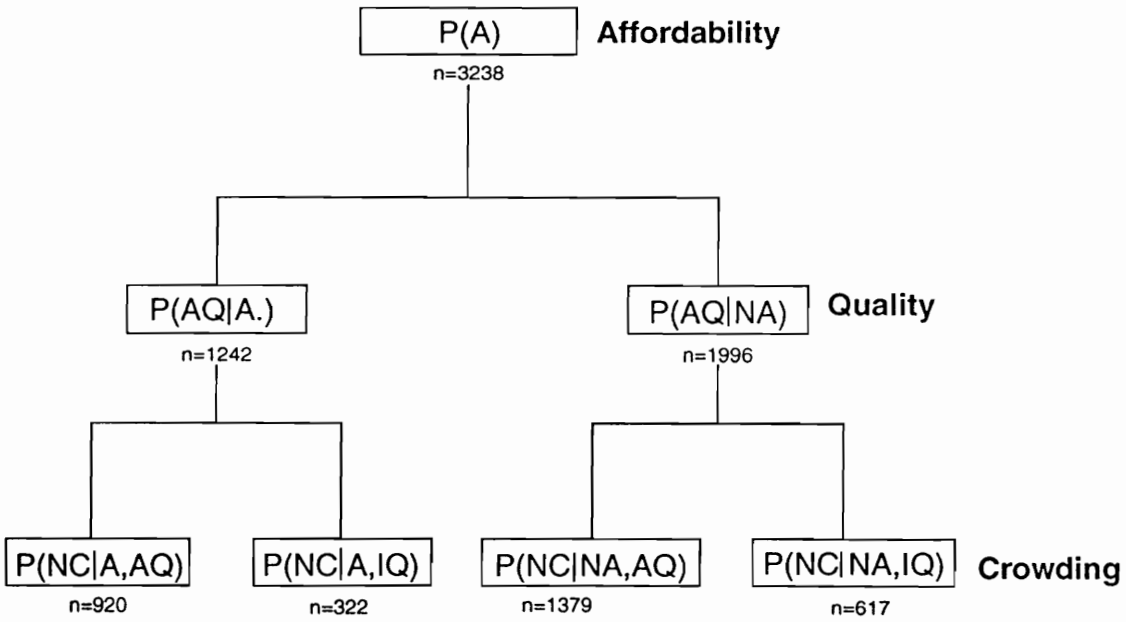


Figure 5.1 Logit Models

adequate quality, uncrowded housing unit. The overall strategy was to estimate these probabilities and subsequently to relate them to the definition of housing deprivation.

The affordability model predicts the likelihood that a household occupies affordable housing. This model contains fifteen independent variables plus an alternative specific constant (CONSTANT). The next two models estimate the probability of adequate quality housing given the affordability choice. Since housing quality generally declines as a property ages, these models include two dummy variables (UNIT AGE < 10 and UNIT AGE < 20) indicating the relative age of the housing unit. At the final level of the housing choice hierarchy are four crowding models. These models are conditioned on prior choices of affordability and quality.

The logit analysis provides estimates of the independent contributions of various demographic and socioeconomic characteristics on the likelihood of **not** having housing problems. With the exception of total household income and the housing assistance variables, these variables represent characteristics of the household head or reference person. The models also include variables which describe the location and characteristics of the housing unit. The positive and negative signs associated with the parameters of these independent variables indicate whether the variable increases or decreases an associated probability. Asymptotic *t* statistics determine if the variable's parameter estimate is significantly different form zero. The following sections discuss each of the estimated models in detail.

5.3.1 Relationship Between Independent Variables and the Affordability Choice

This model predicts the probability that a low-income household occupies

affordable housing where affordability is defined as housing costs which do not exceed thirty percent of income. Consider first the income related variable, INCOME. The accompanying hypothesis is that the probability of finding affordable housing increases with income. Exactly how much income a given household allocates to housing depends both on the availability of relatively low cost housing in an area, and on the households preference for housing relative to other goods (Venti and Wise 1984). However, as household income increases, the supply of very low cost housing becomes less important and households are more likely to find affordable housing.

Two variables in the model reduce the effective price of housing and thus make housing more affordable. These are federal housing subsidies, FEDERALLY ASSISTED, and local housing subsidies, LOCALLY ASSISTED. The AHS data does not indicate the exact nature of or amount of housing assistance. However, the recipient of housing assistance, in most instances, must spend thirty percent of adjusted income on housing, or ten percent of gross income, or the assistance is specific to housing costs (Newman and Schnare 1988). Housing costs which consume more than thirty percent of income are possible. Despite this, most recipients of assistance should live in affordable housing and a positive sign is anticipated.

Including demographic variables enhance the explanatory power of the model. Demographic characteristics affect preferences. Several demographic variables should affect affordability. Blacks experience discrimination in the rental housing market (see for example Galster 1987). It is reasonable to argue that whites can more easily locate housing which consumes less than thirty percent of their household income. The expectation is a positive sign for RACE. Age affects

earnings. As a person ages, income first increases then declines. In tenure choice models, age is generally found to be a positive and significant predictor of ownership (Börsch-Supan and Pollakowski 1986, Rosen 1985, Kain and Quigley 1972). However, when considering low-income renters, age may reflect a declining ability to earn a sufficient amount of money to meet needs. To the extent that this is true, age would have a negative impact on the probability of finding affordable housing. Since women generally earn less than men, female headed households (SEX = 1) should have more difficulty finding affordable housing. Here, a negative sign is expected. As the number of persons in the household increases, households may find it necessary to rent a larger and probably more costly unit. Thus, the variable for household size, PERSONS, should carry a negative sign. Finally, the variable indicating a married household head, MARRIED, is expected to carry a positive sign in the affordability sub-model. Since many married couples both work, the total household income should be higher and the rent burden lower.

Housing costs in the west, the base case, and in the north east are higher than in other parts of the country (Urban Land Institute 1992). The signs associated with the north central and southern parts of the country should be positive, indicating relatively greater affordability. The sign of NORTH EAST should be positive if housing costs there are less than those in the west and negative if more.

Results of the Affordability Model. The affordability model has a rho-square (pseudo R^2) of 0.3118. This is well within the boundaries of from 0.2 to 0.4 which characterize a well fitting model (McFadden 1981). Table 5.3 contains the estimated logit coefficients, the associated standard errors and supporting statistics. These coefficients indicate the incremental change in the odds of finding affordable housing related to a unit change in the independent variable. For all sixteen

Table 5.3

AFFORDABILITY: Probability of Affordable Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	-4.0225	0.2875	13.9913
INCOME	0.2444	0.0089	27.3526
RACE	-0.0785	0.1160	0.6771
SEX	-0.3005	0.1023	2.9382
PERSONS	-0.2544	0.0382	6.6570
MARRIED	-0.1429	0.1236	1.1562
AGE	-0.0123	0.0036	3.9295
GRADE < = 8th	0.8220	0.1921	4.2780
GRADE < = 12th	0.3799	0.1021	3.7203
FEDERAL ASSIST	2.4640	0.2180	11.3017
LOCAL ASSIST	1.6534	0.3561	4.6426
NORTH EAST	-0.3921	0.1557	2.5184
NORTH CENTRAL	1.0830	0.1332	8.1306
SOUTH	1.0695	0.1239	8.6332
MULTI-FAMILY	0.0138	0.1081	0.1277
CENTRAL CITY	-0.0179	0.0929	0.1924

-2 (Log Likelihood) = 1344.568 (p=0.0001)
Pseudo R^2 = 0.3118
Number Of Observations = 3238

variables, estimated coefficients are collectively significantly different from zero as evidenced by the statistic -2 Log Likelihood, that is, $-2\ [log\ L(\beta_0) - log\ L(\beta_k)]$. This statistic is distributed asymptotically chi-square with k-1 degrees of freedom. The value of 1344.568 has an associated probability of less than one percent. Regarding the statistical significance of individual variables, the appropriate statistic is the asymptotic *'t'*. Under the null hypothesis, β_i equals zero, a calculated asymptotic *t-value* exceeding plus or minus 1.96 leads to rejection at the five percent level of significance using a two-tailed test (Wrigley 1985). Rejection at the one percent level requires a *t-value* exceeding 2.576.

As expected, income is a significant predictor of finding affordable housing. The calculated *t* is 27.3526 which leads to rejecting the null hypothesis even at the one percent level of significance. The housing assistance variables, FEDERALLY ASSISTED and LOCALLY ASSISTED, are also significant. Although the *t-values* are smaller, the null hypothesis is rejected at the one percent level. There are a number of other variables which are significant at the one percent level: SEX, PERSONS, AGE, GRADE <=8, GRADE <=12, NORTH CENTRAL, and SOUTH. The variable NORTH EAST is significant at the five percent level and four variables, RACE, MARRIED, MULTI-FAMILY and CENTRAL CITY are not significant.

The sign associated with the SEX variable is negative. This is as expected. Female headed households are less likely to live in affordable housing. In addition to reflecting the generally lower earning power of women, this result may reflect discrimination in rental markets. Landlords may be reluctant to rent to women with children or to women who depend on AFDC as a source of income. Similarly, PERSONS and AGE carry negative signs. Larger households and elderly

households have a lower probability of finding affordable housing. In spite of the fact that rental assistance has gone disproportionately to older renters, they still have a lower probability of obtaining affordable housing.

The fact that the race of the head (RACE) is not significant suggests that there is no difference between whites and non-whites in their ability to locate affordable housing. However, the lack of significance may not be due to race per se but to either the low number of non-whites in the sample (22 percent) or that non-whites are prevented from living in areas where rents would consume more of their income (Galster 1987).

Next, compare the regional variables in both sign and level of significance. The northeastern section of the country is negatively associated with the probability of finding affordable housing at a five percent level of significance while both the north central and southern regions are strongly positive. These variables perform as expected, more affordability problems are found on either coast than in the rest of the country. The insignificance of both multi-family (MULTI-FAMILY) and relative location in the urban area (CENTRAL CITY) variables indicates that neither variable has an effect on the probability of finding affordable housing.

5.3.2 Relationship Between Independent Variables and the Quality Choice

As modeled here, households are expected to choose between adequate quality housing and inadequate quality housing at the second level of the housing choice hierarchy. The quality sub-model is conditioned on the actual level of affordability. The estimated parameters of the resulting two sub-models, their associated standard errors and test statistics are found in Tables 5.4 and 5.5.

Table 5.4

QUALITY - Probability of an Adequate Quality, Affordable Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	-0.6740	0.4779	1.4104
INCOME	0.0262	0.0145	1.8001
RACE	0.1436	0.1795	0.8000
SEX	0.0440	0.1561	0.2821
PERSONS	0.0015	0.0563	0.0894
MARRIED	-0.3377	0.1817	1.8585
AGE	-0.0034	0.0058	0.5924
GRADE < = 8th	-0.6278	0.2773	2.2639
GRADE < = 12th	-0.1331	0.1560	0.8531
FEDERAL ASSIST	0.4044	0.3550	1.1390
LOCAL ASSIST	0.2703	0.5901	0.4580
NORTH EAST	0.9241	0.2574	3.5902
NORTH CENTRAL	0.8681	0.1993	4.3557
SOUTH	0.3231	0.1877	1.7217
MULTI-FAMILY	0.5041	0.1585	3.1808
CENTRAL CITY	-0.2217	0.1435	1.5445
UNIT AGE <10	1.6035	0.2299	6.9761
UNIT AGE < 20	0.9961	0.1603	6.2134

-2 (Log Likelihood) = 137.970 (p=0.0001)
Pseudo R^2 = 0.0971
Number Of Observations = 1242

Table 5.5

QUALITY - Probability of an Adequate Quality, Non-Affordable Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	-1.9636	0.3093	6.3476
INCOME	0.0376	0.0091	4.1533
RACE	0.5494	0.1247	4.4069
SEX	0.1963	0.1207	1.6262
PERSONS	-0.0185	0.0386	0.4800
MARRIED	0.2074	0.1489	1.3926
AGE	0.0065	0.0039	1.6451
GRADE < = 8th	-0.5154	0.2001	2.5758
GRADE < = 12th	-0.2528	0.1230	2.0557
FEDERAL ASSIST	0.7793	0.3076	2.5338
LOCAL ASSIST	-0.2522	0.3452	0.7304
NORTH EAST	1.5608	0.1844	8.4653
NORTH CENTRAL	1.5600	0.1669	9.3488
SOUTH	0.7805	0.1370	5.6956
MULTI-FAMILY	0.2623	0.1236	2.1220
CENTRAL CITY	0.0030	0.1114	0.0205
UNIT AGE <10	1.9834	0.1854	10.6965
UNIT AGE < 20	1.0290	0.1298	7.9290

-2 (Log Likelihood) = 322.944 (p=0.0001)
Pseudo R^2 = 0.1308
Number Of Observations = 1996

The expected signs for the first sixteen variables are the same as those expected in the affordability model. Specification of the quality choice sub-model included, in addition to the original sixteen variables, two variables specifically designed to capture market effects. These variables are two dummy variables, UNIT AGE < 10 and UNIT AGE < 20, classifying the age of the housing unit. Here, the base case is UNIT AGE > 20 or the oldest units. Since quality declines as a unit ages, any units newer than those in the base case should cause an increase (a positive sign) in the probability of living in an adequate quality dwelling.

Results of the Quality Model. The first quality sub-model, conditioned on occupying affordable housing, estimates the probability of adequate quality given that the unit is affordable. The second quality sub-model estimates the probability of occupying adequate quality given that the unit is not affordable. The rho-square of the first sub-model (Table 5.4), where the housing unit is affordable, is .0971 and of the second sub-model (Table 5.5), where housing is not affordable, is .1326. Neither sub-model explains the quality choice very well although the second sub-model does slightly better than the first. A possible reason for the low explanatory power may be difficulty in accurately measuring and categorizing quality using a simple binary variable. A number of transformations of the independent variables, used in an attempt to improve the model fit, did not produce substantially different results. Another explanation for the quality model's low overall explanatory power is that resident's perception of quality may differ from that of policy analysts. For example, residents may not be particularly sensitive to peeling paint or damp basements. Alternatively, low-income households may not perceive the availability of quality choices in the market or they may not search for better quality units. In the absence of a better indicator, the necessary assumption is that quality choices are

not explained very well by the set of independent variables used in the model.

In spite of difficulties in modeling quality choice, a number of points about the parameters of the models are relevant. First, consider the quality sub-model conditioned on finding an affordable housing unit. As is apparent in Table 5.4, the only significant demographic variable is GRADE $<=8$ and no socioeconomic variables are significant. The education level variable is negative suggesting that this group is willing to trade quality for affordability. There are also significant regional variations. Living in the northeast and the north central sections of the country increase the probability of finding adequate quality housing with *t-values* of 3.5971 and 4.3422 respectively. The variable which identifies whether the unit is multi-family (MULTI-FAMILY) is positive and significant. This indicates an increase in the probability of finding adequate quality when the household occupies a multi-family unit over a single family unit. It appears that households sacrifice quality for the opportunity to live in a single family dwelling. The location variable (CENTRAL CITY) is not significant. Evidently, location does not affect the household's ability to find housing of adequate (or inadequate) quality when rent consumes less than thirty percent of income. However, the variables which indicate the approximate age of the structure (UNIT AGE $<$ 10 and UNIT AGE $<$ 20) are both positive and significant at the one percent level. The estimated parameter for UNIT AGE $<$ 10 is larger than the estimated parameter for UNIT AGE $<$ 20 confirming expectations that newer housing units are of better quality or at least in better repair.

The results obtained with the second quality sub-model, Table 5.5, are somewhat different from those discussed above. These households do not occupy affordable housing and more independent variables are significant predictors of the

adequate quality choice. The income variable (INCOME) is significant at the one percent level although the parameter estimate is relatively small, .039. This variable is income expressed in thousands, therefore, an additional thousand dollars of income will only increase the probability of adequate quality from 21.43 percent to 22.09 percent[6]. Federal housing assistance is significant but local housing assistance is not. This result may reflect the low incidence of local housing assistance. The race of the head of the household (RACE) is positive and significant. This outcome suggests that white households are more likely than non-white households to find adequate quality housing when they spend more than thirty percent of their income on housing. Alternatively, even if a non-white household spends more than thirty percent of income on housing, there is a lower probability of finding a housing unit of adequate quality. If non-whites do not have a different perception of quality, the outcome may reflect discrimination in housing markets. Both education variables (GRADE $<=8$ and GRADE $<=12$) are negative and significant at the five percent level. The base case is GRADE >12 which describes the highest level of education. Individuals with lower levels of education have a lower probability of finding adequate quality housing. These households may lack knowledge about how to either locate or evaluate housing of adequate quality.

Two regional variables, NORTH CENTRAL and SOUTH, are positive and significant with NORTH CENTRAL having the higher parameter estimate. Housing of lower quality is more prevalent on either coast and, for households spending more than thirty percent of income on housing, the probability of adequate quality is greatest in the midwest. This result seems to refute the common perception of poor quality southern housing. However, recall that the analysis

[6] As noted earlier, this probability is estimated at the mean level of the continuous variables and the zero level of the dummy variables.

covers only urban renters. A large proportion of the low-income southern population is rural. Finally, consider the variables related to the housing unit, MULTI-FAMILY, UNIT AGE < 10 and UNIT AGE < 20. Each of these variables is positive and significant. MULTI-FAMILY is significant at the five percent level and the variables related to the age of the unit are significant at the one percent level. The association of multi-family units with better quality may be due to the fact that many apartment owners are large corporations which have property management departments overseeing day-to-day operations. In contrast, the single family rental may be owned by an individual with only a few rental properties where maintenance is more sporadic or not performed at all.

In summarizing the results obtained from the adequate quality sub-models, there are two important findings. The finding of greatest interest is that quality choice is not easily modeled as a binary decision. This finding suggests that conclusions drawn by various researchers (Apgar 1992, Nelson and Khadduri 1992, Newman and Schnare 1988) concerning the incidence of inadequate quality housing for particular groups may not accurately reflect the choices made by households or their living conditions. Using the modified HUD index of housing quality with AHS data may not be an adequate tool for formulating housing policy. The second finding is related to the differences found between the sub-model conditioned on living in affordable housing and the sub-model conditioned on living in non-affordable housing. The importance of a number of independent variables differs between these two models. Finding adequate quality housing is more closely tied to socioeconomic and demographic variables when households spend more than thirty percent of income on housing than when households spend less. We know that housing is idiosyncratic both physically and in the way it is marketed. Households

spending less than thirty percent of income on housing may not be identifiable by their socioeconomic and demographic characteristics because they are in some of these idiosyncratic situations. That is, they may not really be paying market rent or they found other discontinuities in the housing market.

5.3.3 Relationship Between Independent Variables and the Crowding Choice

The crowding models, estimated at the third level of the housing choice hierarchy, are conditioned on actual choices made at the previous two levels. These models predict the probability of no crowding or that the persons per room (PPR) indicator does not exceed one. The set of independent variables in the models include all sixteen variables used in the affordability model.

Results of the Crowding Models. Tables 5.6 through 5.9 present the results of the four models estimated. In the total sample of 3238 households, there are only 212 (.0655 percent) crowded households. Since these models are conditioned on both affordability and quality choices, the incidence of crowding in any of the sub-models is quite small[7]. All of the models preform well as measured by the pseudo R^2.

The two variables which strongly and consistently influence the probability of no crowding in all four sub-models are PERSONS and MULTI-FAMILY. Both are negative and significant. Indeed, the high *t-values* indicate that these variables do most to explain the incidence of crowding. This leads to questions about how redefining crowding (i.e. raising the persons-per-room indicator to say 1.5) would

[7] To ascertain that the models were accurately reflecting the probability of no crowding and not driven by the low incidence of crowding in the data, I estimated the models using a subset of the data which included all of the crowded households and an equal number of non-crowded households. The results obtained from these additional models were substantially the same as obtained using the full data set.

Table 5.6

CROWDING: Probability of Not Crowded - Affordable & Adequate Quality Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	6.2289	2.0877	2.9837
INCOME	0.0177	0.0599	0.2956
RACE	0.3554	0.6552	0.5424
SEX	0.6416	0.6627	0.9682
PERSONS	-1.6705	0.2308	7.2368
MARRIED	0.5298	0.6127	0.8648
AGE	0.0809	0.0371	2.1820
GRADE < = 8th	-2.2425	0.9230	2.4295
GRADE < = 12th	-0.4048	0.6174	0.6556
FEDERAL ASSIST	0.9536	1.3127	0.7265
LOCAL ASSIST	0.8464	1.3894	0.6093
NORTH EAST	1.6168	0.8806	1.8361
NORTH CENTRAL	1.4038	0.6788	2.0679
SOUTH	1.9560	0.7580	2.5806
MULTI-FAMILY	-2.0632	0.6931	2.9769
CENTRAL CITY	0.4061	0.5101	0.7960

-2 (Log Likelihood) = 170.627 (p=0.001)
Pseudo R^2 = 0.5643
Number Of Observations = 920

Table 5.7

CROWDING: Probability of Not Crowded - Affordable, Inadequate Quality Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	7.9306	3.3305	2.3821
INCOME	0.1627	0.1007	1.6163
RACE	1.8518	1.0302	1.7975
SEX	3.1402	1.2606	2.4909
PERSONS	-2.8495	0.6266	4.5413
MARRIED	-0.5336	1.0510	0.5077
AGE	0.0669	0.0485	1.3780
GRADE < = 8th	-1.2128	1.3539	0.8958
GRADE < = 12th	0.9583	1.2733	0.7526
FEDERAL ASSIST	-1.4591	1.8414	0.7924
LOCAL ASSIST	-3.4184	3.2380	1.0557
NORTH EAST	4.6993	3.0756	1.5279
NORTH CENTRAL	1.2297	1.0148	1.2117
SOUTH	1.7682	1.1470	1.5417
MULTI-FAMILY	-3.5965	1.1630	3.0924
CENTRAL CITY	1.3050	0.8154	1.6005

-2 (Log Likelihood) = 150.927 (p=0.001)
Pseudo R^2 = 0.7423
Number Of Observations = 322

Table 5.8

CROWDING: Probability of Not Crowded - Non-affordable, Adequate Quality Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	14.0116	2.0005	7.0039
INCOME	0.0491	0.0333	1.4714
RACE	-0.2453	0.5475	0.4481
SEX	-1.0229	0.5573	1.8353
PERSONS	-2.5596	0.3039	8.4235
MARRIED	-0.4237	0.5677	0.7463
AGE	0.0757	0.0260	2.9154
GRADE < = 8th	-3.0164	0.7207	4.1850
GRADE < = 12th	-0.7421	0.5901	1.2577
FEDERAL ASSIST	2.5521	1.0232	2.4941
LOCAL ASSIST	3.5575	1.6335	2.1779
NORTH EAST	0.0822	0.6361	0.1292
NORTH CENTRAL	1.4252	0.6495	2.1945
SOUTH	0.2978	0.5494	0.5420
MULTI-FAMILY	-2.8787	0.6136	4.6914
CENTRAL CITY	-0.5753	0.4464	1.2886

-2 (Log Likelihood) = 382.175 (p=0.0001)
Pseudo R^2 = 0.6988
Number Of Observations = 1379

Table 5.9

CROWDING: Probability of Not Crowded - Non-affordable, Inadequate Quality Unit

Variable	Parameter	Standard Error	t-value
CONSTANT	7.5941	1.2189	6.2301
INCOME	0.0597	0.0338	1.7628
RACE	-0.3831	0.4148	0.9236
SEX	0.1875	0.4335	0.4324
PERSONS	-1.3876	0.1501	9.2450
MARRIED	0.5511	0.4509	1.2222
AGE	-0.0103	0.0182	0.5668
GRADE < = 8th	-1.2501	0.6062	2.0620
GRADE < = 12th	-0.4745	0.5333	0.8898
FEDERAL ASSIST*			
LOCAL ASSIST	2.5418	1.8881	1.3462
NORTH EAST	1.4058	0.7176	1.9590
NORTH CENTRAL	0.4515	0.5518	0.8182
SOUTH	0.4501	0.4529	0.9938
MULTI-FAMILY	-1.2761	0.4076	3.1306
CENTRAL CITY	0.1064	0.3744	0.2841

-2 (Log Likelihood) = 228.427 (p=0.0001)
Pseudo R^2 = 0.4982
Number Of Observations = 617
*(Note: Parameter Estimate Is Infinite)

affect the probability of no crowding. Also, households living in multi-family structures are more likely to find crowding a problem irrespective of affordability or quality.

The education variables are also negative in almost every model. However, only the variable GRADE < =8 is significant and then in only three of the models. It is not significant in the model which predicts no crowding given that the unit is affordable but of inadequate quality (Table 5.7). We saw earlier that households where the head has little education seem to trade off quality for affordability. The significance of GRADE < =8 suggests that households where the head has less than a high-school education are also more likely to be crowded. As might be expected, AGE, when it is significant, has a positive influence on the probability of no crowding. Young families are just more likely to be crowded. The regional variables are positive, indicating that the probability of not being crowded is smaller in the western part of the country, the base case. These variables are not, however, significant in most of the models. No variables, with the exception of PERSONS and MULTI-FAMILY, are consistently significant.

5.4 Choice of the Housing Bundle

The statistical approach employed in this research provides estimates of marginal (affordability) and conditional probabilities (quality and crowding). This type of model is appropriate because my concern is for the combined choice that an urban household makes given that their income is below eighty percent of regional median income and that they rent (not purchase) housing. Moreover, I am not concerned with the physical housing unit *per se* but with a housing bundle described by the cost of the unit, the quality of the unit, and the size of the unit relative to the

size of the household. Up to this point, the discussion has centered on each of the sub-models independently. The following discussion emphasizes the combined choice.

There are eight possible combinations of affordability, quality and crowding facing a household. These combinations are easily traced in the preference trees illustrated in Figure 5.2. For example, the first combination may represent a household finding affordable, adequate quality, uncrowded housing (Figure 5.2a). A possible second combination is affordable, adequate quality but crowded housing (Figure 5.2b). Subsequent combinations reflect one change at a time until reaching the last combination or the choice of non-affordable, inadequate quality, crowded housing (Figure 5.2c). The probability associated with each of these combined choices is itself a combination of the probabilities of each of the individual choices[8].

The combined probability of each of the eight possible housing choices is easily determined. The equation used to calculate each probability for the individual binary logit equations is:

$$logit(p) = CONSTANT + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_k x_k$$

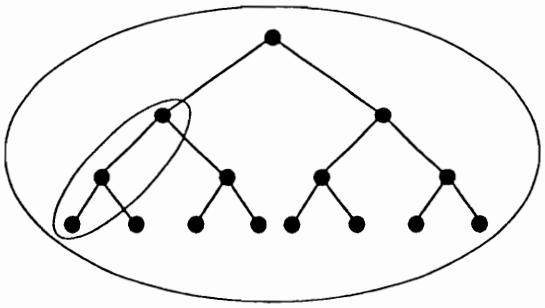
where $CONSTANT$ = the alternative specific constant, β = the parameter estimate, and the probability is given by:

$$P_i = e^{\beta' x_i} / [1 + e^{\beta' x_i}]$$

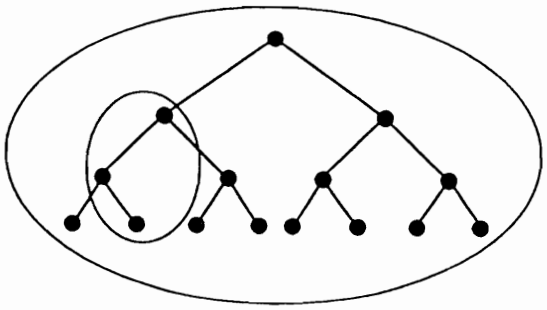
The joint probabilities are combinations of the marginal and conditional probabilities. For each example, the following probabilities form joint probabilities:

$P_x(A)$ = the probability that a household with vector of characteristics x find affordable housing.

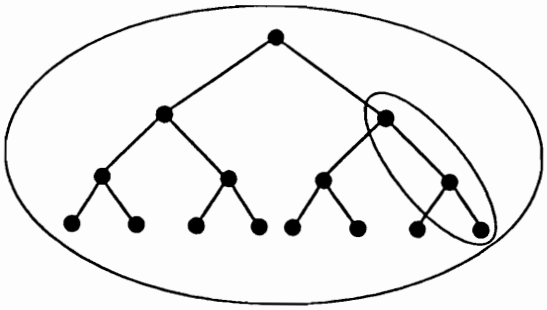
[8] Refer to Section 3.6 for the method.



a. Affordable, Adequate Quality, Uncrowded Housing



b. Affordable, Adequate Quality, Crowded Housing



c. Non-affordable, Inadequate, Crowded Housing

Figure 5.2 Housing Choice Bundle

$P_x(AQ \mid A)$ = the probability that a household finds a unit of adequate quality, given that they live in affordable housing.

$P_x(NC \mid A,AQ)$ = the probability that a household chooses an uncrowded unit, given that they live in affordable, adequate quality housing.

The probability of **not** fitting the definition of housing deprivation, or the probability of basic housing (BH) is determined as follows:

$$P(BH) = P_x(A) \bullet P_x(AQ \mid A) \bullet P_x(NC \mid A,AQ)$$
$$= P(A:AQ:NC)$$

This value can be calculated for every household in the sample. Doing this presents a practical problem since the initial sample is large. To provide additional information about the housing choices of particular population groups, I calculated joint probabilities for three different household types. Nelson and Khadduri (1992) suggest that very low-income families are in most need of housing assistance. My data set is not arranged so that families, versus households consisting of unrelated persons, are identifiable. However, I identify three types of low-income households which reflect typical households. These are a married couple household with a male head; an unmarried, female headed household; and an elderly, unmarried, female headed household. Application of the joint model permits forecasting changes in housing behavior resulting from the receipt of housing assistance. For each household, calculating the marginal probability of affordability and the conditional probabilities of quality and crowding leads to the joint probability of a particular housing bundle.

Tables 5.10, 5.11 and 5.12 present joint probabilities for the three different household types. With the exception of INCOME, SEX, AGE, MARRIED, PERSONS and FEDERALLY ASSISTED; the independent variables are the same

Table 5.10

HOUSING PROBABILITIES by HOUSEHOLD TYPE and REGION

Married, male head, income 35% of median, four person household				
	North East	North Central	South	West
No Assistance				
P(A:AQ:NC)	0.0712	0.1691	0.1151	0.0467
P(NA:AQ:NC)	0.6320	0.5524	0.4201	0.3029
P(A:IQ:NC)	0.0142	0.0256	0.0330	0.0122
P(NA:IQ:NC)	0.2183	0.1678	0.2992	0.4067
P(A:AQ:OC)	0.0047	0.0144	0.0048	0.0098
P(NA:AQ:OC)	0.0377	0.0093	0.0236	0.0209
P(A:IQ:OC)	0.0002	0.0128	0.0122	0.0196
P(NA:IQ:OC)	0.0210	0.0464	0.0903	0.1727
With Federal Assistance				
P(A:AQ:NC)	0.4668	0.6540	0.5504	0.3651
P(NA:AQ:NC)	0.3928	0.1935	0.2116	0.2417
P(A:IQ:NC)	0.0572	0.0297	0.0538	0.0173
P(NA:IQ:NC)	0.0590	0.0266	0.0657	0.1400
P(A:AQ:OC)	0.0132	0.0237	0.0108	0.0402
P(NA:AQ:OC)	0.0018	0.0003	0.0009	0.0013
P(A:IQ:OC)	0.0029	0.0641	0.0856	0.1199
P(NA:IQ:OC)	0.0057	0.0079	0.0198	0.0594

Married, male head, income 80% of median, four person household				
	North East	North Central	South	West
No Assistance				
P(A:AQ:NC)	0.7502	0.7600	0.6472	0.4936
P(NA:AQ:NC)	0.0948	0.0585	0.0948	0.0762
P(A:IQ:NC)	0.0931	0.1081	0.1685	0.1835
P(NA:IQ:NC)	0.0172	0.0113	0.0451	0.0661
P(A:AQ:OC)	0.0394	0.0526	0.0240	0.0974
P(NA:AQ:OC)	0.0024	0.0005	0.0027	0.0024
P(A:IQ:OC)	0.0001	0.0043	0.0067	0.0217
P(NA:IQ:OC)	0.0006	0.0012	0.0060	0.0108
With Federal Assistance				
P(A:AQ:NC)	0.8978	0.8847	0.8258	0.7280
P(NA:AQ:NC)	0.0101	0.0059	0.0118	0.0106
P(A:IQ:NC)	0.0718	0.0714	0.1241	0.1151
P(NA:IQ:NC)	0.0008	0.0005	0.0025	0.0041
P(A:AQ:OC)	0.0191	0.0250	0.0135	0.0695
P(NA:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(A:IQ:OC)	0.0002	0.0122	0.0212	0.0585
P(NA:IQ:OC)	0.0000	0.0001	0.0003	0.0007

Table 5.11

HOUSING PROBABILITIES by HOUSEHOLD TYPE and REGION

Unmarried, female head, income 35% of median, three person household,				
	North East	North Central	South	West
No Assistance				
P(A:AQ:NC)	0.0870	0.2079	0.1428	0.0743
P(NA:AQ:NC)	0.6593	0.5491	0.4344	0.3191
P(A:IQ:NC)	0.0113	0.0299	0.0366	0.0264
P(NA:IQ:NC)	0.2280	0.1895	0.3441	0.4965
P(A:AQ:OC)	0.0009	0.0028	0.0009	0.0024
P(NA:AQ:OC)	0.0055	0.0013	0.0034	0.0031
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0001
P(NA:IQ:OC)	0.0079	0.0188	0.0373	0.0757
With Federal Assistance				
P(A:AQ:NC)	0.5175	0.7162	0.6149	0.4635
P(NA:AQ:NC)	0.3743	0.1798	0.1987	0.2306
P(A:IQ:NC)	0.0445	0.0679	0.1043	0.1049
P(NA:IQ:NC)	0.0589	0.0284	0.0717	0.1631
P(A:AQ:OC)	0.0024	0.0042	0.0019	0.0080
P(NA:AQ:OC)	0.0002	0.0000	0.0001	0.0002
P(A:IQ:OC)	0.0000	0.0002	0.0002	0.0011
P(NA:IQ :OC)	0.0020	0.0028	0.0078	0.0249

Unmarried, female head, income 80% of median, three person household,				
	North East	North Central	South	West
No Assistance				
P(A:AQ:NC)	0.8203	0.8437	0.7279	0.6695
P(NA:AQ:NC)	0.0889	0.0539	0.0896	0.0723
P(A:IQ:NC)	0.0666	0.0805	0.1299	0.1542
P(NA:IQ:NC)	0.0160	0.0109	0.0446	0.0666
P(A:AQ:OC)	0.0070	0.0095	0.0043	0.0206
P(NA:AQ:OC)	0.0003	0.0001	0.0004	0.0003
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:IQ:OC)	0.0002	0.0004	0.0021	0.0039
With Federal Assistance				
P(A:AQ:NC)	0.9363	0.9308	0.8798	0.8439
P(NA:AQ:NC)	0.0092	0.0054	0.0107	0.0097
P(A:IQ:NC)	0.0504	0.0588	0.1043	0.1258
P(NA:IQ:NC)	0.0008	0.0005	0.0024	0.0041
P(A:AQ:OC)	0.0033	0.0043	0.0023	0.0128
P(NA:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0001
P(NA:IQ:OC)	0.0000	0.0000	0.0000	0.0002

Table 5.12

HOUSING PROBABILITIES by HOUSEHOLD TYPE and REGION

Unmarried, elderly female head, income 35% of median, one person household				
	North East	North Central	South	West
No Assistance				
P(A:AQ:NC)	0.0975	0.2292	0.1563	0.0862
P(NA:AQ:NC)	0.6943	0.5656	0.4738	0.3672
P(A:IQ:NC)	0.0199	0.0362	0.0442	0.0321
P(NA:IQ:NC)	0.1937	0.1675	0.3227	0.5078
P(A:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:IQ:OC)	0.0006	0.0014	0.0030	0.0066
With Federal Assistance				
P(A:AQ:NC)	0.5438	0.7322	0.6280	0.4900
P(NA:AQ:NC)	0.3583	0.1676	0.1926	0.2362
P(A:IQ:NC)	0.0519	0.0772	0.1186	0.1219
P(NA:IQ:NC)	0.0459	0.0228	0.0602	0.1458
P(A:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:IQ:OC)	0.0001	0.0002	0.0006	0.0020

Unmarried, elderly female head, income 80% of median, one person household				
No Assistance	North East	North Central	South	West
P(A:AQ:NC)	0.8314	0.8524	0.7329	0.7004
P(NA:AQ:NC)	0.0822	0.0497	0.0867	0.0725
P(A:IQ:NC)	0.0746	0.0896	0.1448	0.1714
P(NA:IQ:NC)	0.0118	0.0082	0.0354	0.0553
P(A:AQ:OC)	0.0000	0.0000	0.0000	0.0001
P(NA:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:IQ:OC)	0.0000	0.0000	0.0001	0.0003
With Federal Assistance				
P(A:AQ:NC)	0.9352	0.9296	0.8731	0.8489
P(NA:AQ:NC)	0.0082	0.0048	0.0099	0.0091
P(A:IQ:NC)	0.0746	0.0652	0.1151	0.1387
P(NA:IQ:NC)	0.0118	0.0004	0.0018	0.0032
P(A:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:AQ:OC)	0.0000	0.0000	0.0000	0.0000
P(A:IQ:OC)	0.0000	0.0000	0.0000	0.0000
P(NA:IQ:OC)	0.0000	0.0000	0.0000	0.0000

for all three household types. The values of those variables are; RACE equals one (a white household head), GRADE < =8 is zero, GRADE < =12 is one, MULTI-FAMILY is one (indicating a multi-family unit), and CENTRAL CITY is zero (a suburban unit). The UNIT AGE variable, used in the quality sub-model, indicates a unit of between ten and twenty years old[9]. These income level of the hypothetical households is a percentage of the median income for their particular region. The mean income level for each sub-model, Table 5.13, provides additional insight into the role income plays in the probability of a particular housing combination.

There are a number of interesting similarities between the tables. With all three household types, the greatest problems are related to affordability when the household income is at thirty-five percent of regional median and the household does not receive assistance. At that income level, the probability of living in non-affordable, adequate quality, uncrowded housing, *P(NA:AQ:NC)*, ranks first for every region but the west. The combination with the second highest probability is the probability of non-affordable, inadequate quality, uncrowded housing, *P(NA:IQ:NC)*. In the western region, the position of those two joint probabilities is reversed. This indicates that households located in the western region have a more difficult time finding housing of adequate quality even if they spend more than thirty percent of income on housing. At the eighty percent income level for unassisted households, the highest probability is affordable, adequate quality, uncrowded housing, *P(A:AQ:NC)*. These are the households with no problems.

For the married household living in the northeast, earning eighty percent of median income (Table 5.10), combining the probability of affordable housing with the probability of adequate quality and not crowded, *P(A:AQ:NC)*, results in a joint

[9] To understand how changing these variables affects the associated probabilities refer to Table 5.13.

Table 5.13

Signs and Significance of Independent Variables by Sub-Model

Variable	P(A)	P(AQ|A)	P(AQ|NA)	P(NC|A:AQ)	P(NC|A:IQ)	P(NC|NA:AQ)	P(NC|NA:IQ)
CONSTANT	- **	- *	- **	+ **	+ *	+ **	+ **
INCOME	+ **	+	+ **	+	+	+	+
RACE (1 = white)	-	+	+ **	+	+	-	-
SEX (1 = female)	- **	+	+	+	+ *	-	+
PERSONS	- **	+	-	- **	- **	- **	- **
MARRIED (1 = yes)	-	+	+	+	-	-	-
AGE	- **	-	+	+ *	+	+ **	-
GRADE <=8 (1 = yes)	+ **	- *	- **	- *	-	- **	- *
GRADE <=12 (1 = yes)	+ **	-	- *	-	+	-	-
FEDERAL ASSIST (1 = yes)	+ **	+	+ *	+	-	+ *	**na**
LOCAL ASSIST (1 = yes)	+ **	+	-	+	-	+ *	+
NORTH EAST (1 = yes)	- *	+ **	+ **	+	+	+	+ *
NORTH CENTRAL (1 = yes)	+ **	+ **	+ **	+ *	+	+ *	+
SOUTH (1 = yes)	+ **	+	+ **	+ **	+	+	+
MULTI-FAMILY (1 = yes)	+	+ **	+ *	- **	- **	- **	- **
CENTRAL CITY (1 = yes)	-	-	+	+	-	-	+
UNIT AGE < 10 (1 = yes)	**na**	+ **	+ **	**na**	**na**	**na**	**na**
UNIT AGE < 20 (1 = yes)	**na**	+ **	+ **	**na**	**na**	**na**	**na**

* = Significant at the 5% level
** = Significant at the 1% level

probability of 75.02 percent. That probability increases to 89.78 percent with federal assistance. However, if the households' income is fifty percent[10] of northeastern regional median, the percentages are 24.09 without federal assistance and 73.35 with federal assistance. If the income is thirty-five percent of median, the percentages are 7.12 and 46.68 respectively. Obviously, the benefit of federal housing assistance increases as income level declines. Similar though not directly proportional results occur in other regions of the country and with the other two household types (Tables 5.11 and 5.12).

5.5 Concluding Comments

Development of housing policies to help alleviate the housing problems of low-income households must be based on a clear understanding of what the problem is and what the magnitude of the problem is. Nelson and Khadduri (1992) make a case for targeting housing policy to "worst case" needs or households with housing cost burdens in excess of fifty percent of income, or severely inadequate housing among unassisted renters. As Nelson and Khadduri note; "the likelihood and severity of housing problems vary sharply with income, so that the very poorest households -most of them renters- have by far the most pervasive and serious problems (p. 50)." Arguments that emphasize focusing housing assistance on the very low-income are supported by the results of the empirical analysis presented in this chapter.

Table 5.13 provides a snapshot look at signs and level of significance for variables in the seven sub-models. Some general observations are that the probability of finding affordable housing for someone with education beyond high-

[10] Table 5.9 does not present probabilities for fifty percent of median income.

school is lower than for other education levels but the probability of adequate quality is greater. A household where the head is white has a lower probability of choosing affordable housing but a higher probability of finding adequate quality than does a household where the head is not white. In the models where the household size variable is significant, it is negative. That is, larger families have more severe housing problems. Multi-family housing is positively associated with both affordability and quality but negatively associated with crowding. Finally, with the exception of NORTH EAST in the affordability model, all of the regional variables are positively associated with the outcomes. It appears that households in the west suffer from the most severe housing problems.

The probability of selecting housing of adequate quality differs for households in affordable housing and non-affordable housing. The household living in affordable housing has a higher probability of also living in adequate quality housing. The household living in non-affordable housing has a lower probability of adequate quality. This might seem contradictory at first glance. However, consider the median income of both groups. The mean income level of the group living in affordable housing is $19,359 while for the group in non-affordable housing it is $11,211 (Table 5.14). Obviously, it is the very low-income who have the most severe housing problems. This result is confirmed to some extent in the crowding models. Here, the households occupying affordable housing whether adequate or inadequate quality housing have incomes well above those in non-affordable housing.

The question of which households need assistance is not ultimately a question of household type but a question of income level. Clearly, households whether a married couple, a single mother or an elderly person have the most severe problems if they also have very low-incomes. Federal housing assistance improves but does

Table 5.14

Income Levels for Sub-Models - Mean and Standard Deviation

Model	Mean	Standard Deviation	Sample Size
P(A)	$14,336	$7,527	3238
P(AQ|A)	$19,359	$5,783	1242
P(AQ|NA)	$11,210	$6,757	1996
P(NC|A:AQ)	$19,629	$5,714	920
P(NC|A:IQ)	$18,589	$5,917	322
P(NC|NA:AQ)	$11,781	$6,899	1379
P(NC|NA:IQ)	$ 9,935	$6,246	617

not assure a households chances of finding affordable, adequate quality, uncrowded housing. The final chapter, Chapter 6, summarizes this research and discusses further the outcome of the empirical analysis.

CHAPTER 6
SUMMARY AND CONCLUSIONS

Rent-to-income ratios, quality ranking measures and crowding are frequently used to describe housing problems. Arguments about the inadequacy of these measures rest upon the rigidity of the standard assigned to each measure. For example, a thirty percent rent-to-income ratio may not constitute a burden for a middle-class household but be excessive for a low-income household. Yet, these measures continue to be used. One purpose of this research was to evaluate housing choices using common indicators to predict the joint probability of housing problems for individual households. A second purpose was to judge the value of this type of analysis in designing housing policy. This chapter presents a summary of the research and discusses conclusions based on the empirical analysis. Final comments include limitations of the study and suggestions for further research.

6.1 Summary

Many discussions of housing problems revolve around a simple enumeration of the number of households classified as having affordability, quality or crowding problems. In contrast, the statistical technique used in this research models these

housing problems as choices low-income renters must make in order to secure shelter. The study presents empirical evidence on the likelihood that a low-income, urban, renter household will have multiple housing problems. Specifically, I modeled the probability of affordable, adequate quality, uncrowded housing using seven binary logit models. Chapter 1 addresses the reasons for using discrete choice modeling to examine the housing situation of low-income urban renters. Since housing assistance is not universally available to the low-income household, this research also questions the difference between the housing choices of those low-income renter households who rely solely on income to pay rent and the choices made by households receiving assistance.

Chapter 2 contains a review of literature pertinent to this research. The first part of the chapter reviews the literature on housing standards and indicators. This review establishes a historic basis for the measures of affordability, quality and crowding used in the empirical analysis. Next, research related to the impact of housing policies is considered. There is a lack of consensus on the effect of housing assistance. Many economists argue that housing assistance acts as an income transfer (see Venti and Wise 1983 or Haunshek and Quigley 1981) and thus has only marginal impacts on housing conditions. Others (Hammond 1983, Murray 1980) argue that housing assistance programs do alter consumption.

The final part of the chapter includes a discussion of research using housing choice models. The model developed in this dissertation is an extension of those models. Almost all of the models reviewed examine either tenure choices or housing structure type choices (Börsch-Supan 1990, Börsch-Supan and Pitkin 1988, Clark and Onaka 1985, Boehm 1979, Quigley 1976). Standard practice is to assume a random utility framework and model the housing choice as a discrete rather than a

continuous function. Discrete choice modeling is found to be an appropriate technique because many of the decisions related to securing housing are discrete.

Research methodology is the subject of Chapter 3. This chapter describes the use of qualitative choice analysis to estimate the probability that a particular household occupies affordable, adequate quality, uncrowded housing. The model provides estimates of this probability based on characteristics of the household and housing unit, regional location of the household and the receipt (or non-receipt) of housing assistance. The utility of each choice is a combination of the attributes of the alternatives and the characteristics of the household. Individual households maximize utility within a random utility framework. Random utility maximization assumes that the utility function contains both a deterministic component and a random component (Ben-Akiva and Lerman 1985, McFadden 1974 1978 1981). The researcher observes only the deterministic component. The presence of a random component renders the situation probabilistic.

The discrete choice model specifies the parameters of the deterministic portion of the random utility function. The random components are assumed to follow one of three distributions; a normal distribution, a logistic distribution or an extreme value distribution. The logistic distribution is the most tractable of these and is the one used in this research. Since the purpose of the model is to classify households as making a series of binary decisions, each model is a binary logit. Logit models involve the estimation of a logistic regression equation which classifies observations according to a critical (0,1) value of the dependent variable. Maximum likelihood estimation provides a solution to the logistic regression function.

Critical values of the dependent variables arise from the definitions of housing affordability, quality and crowding. The definition of affordability is a rent-

to-income ratio of thirty percent or less. A HUD related quality index (Newman and Schnare 1988) is the basis for the binary quality definition. Crowding is a persons-per-room indicator. I estimated the affordability model first followed by the quality model and the crowding model. The quality model was conditioned on the housing units actual affordability while the crowding model was conditioned on both actual unit affordability and actual unit quality. The final section of the chapter describes the joint probability function.

The data for this research are from the 1989 American Housing Survey (AHS). Chapter 4 describes this data set and the subsequent sample of low-income urban renters used in the analytical portion of the paper. The sample, 3238 households, includes only low-income, urban households who moved during the previous twelve month period. A recent mover sample better reflects current market conditions. Income levels in the sample are at or below eighty percent of the 1989 regional median income.

Chapter 5 contains the empirical results. The initial section of the chapter presents a description of both the dependent and the independent variables used in the analysis. Following that, individual sub-models (Tables 5.3 through 5.9) focus on the separate problems of affordability, quality and crowding. Combining individual sub-models leads to the estimation of joint probabilities. Within the discrete choice framework, Tables 5.10, 5.11 and 5.12 illustrate the probability distribution of three representative households with respect to finding particular combinations of affordability, quality and crowding. The probability that a household makes a particular housing decision depends heavily on income level.

6.2 Conclusions

In this final chapter an important point needs to be emphasized. It concerns the idiosyncratic nature of housing markets and their role in shaping the decisions household members make about housing. Individual housing markets are largely ignored when using a model like this. Obviously, this omission obscures that relationship. However, neither can patterns of housing problems be explained by market forces alone. This study argues that socioeconomic and demographic variables and government policies do help explain housing outcomes for low-income urban renters. The calculated probabilities serve as a frame of reference which provides cues to persons making housing policy decisions. The following discussion elaborates on this theme by addressing the questions asked in the first chapter. It is clear, however, that a statistical analysis alone cannot account for the enduring housing problems of the low-income population.

6.3 Implications in Terms of the Problem Statement

The purpose of the first five chapters of this dissertation was to set out a model of housing choice which better explains the situation of low-income urban renters than does simple enumeration of data. This research portrays the housing bundle as having three dimensions. In order to evaluate the results of portraying housing in this way, I return to the questions asked in the first chapter.

1. Can a discrete choice model represent housing choices made by low-income renters with regard to affordability, quality and crowding? Will this methodology provide a means of validating the survey-based estimates that dominate the literature on housing problems?

First of all, it is not the case that there are two discrete sets of outcomes for each sub-model described in this research. Indeed, although these models do seem to represent housing choices, it is important to remember that, with the exception of the crowding model, the underlying variables are continuous and attempts to model them as binary variables ignores information that continuous variables might provide. The model assumes that households evaluate their housing as either affordable or non-affordable, adequate quality or inadequate quality, or as not crowded or crowded. By treating these variables as binary variables, it is possible to tie this model to specific policy issues. The results of the estimated models are consistent with previous survey-based estimates in the literature (Nelson and Khadduri 1992, Apgar 1992, Newman and Schnare 1988). The model performs well in identifying the same problems in similar population groups. The discrete choice model is appropriate, both theoretically and operationally, to the evaluation of multiple housing problems. However, the definition of quality and subsequent model should be reevaluated in an attempt to find a model that fits the data better.

2. Can this model be used to document differences between housing choices made by different demographic groups and in different regions of the country?

The basic model used in this paper does seem to provide useful estimates of the current housing behavior of particular demographic groups. The households are distinguished by such demographic characteristics as age, sex, race and marital status. Rather than documenting differences, the model seems to document similarities between groups. That is, a household living in the northeast and earning less that thirty-five percent of regional median income has less than a ten percent chance of finding affordable, adequate quality housing whether the household consists of a married couple, a single mother or an elderly female. The importance

of this result may be better understood when it is contrasted with simply looking at the proportion of households in a particular category who have housing problems. Through a comprehensive tabulation of the 1989 AHS data, Nelson and Khadduri (1992) found that elderly renters had the fewest housing problems. They also point out that the elderly are generally eligible for rental housing programs (p.12). What enumeration of the data does not show is that a single, very low-income, elderly woman has about the same probability of finding basic shelter as do other very low-income demographic groups. There are striking regional contrasts. Households of every type living in the western part of the United States have a much lower probability of finding basic shelter.

The model does draw attention to the tradeoffs some households make between the various components of affordability, quality and crowding. By examining the signs and levels of significance associated with the independent variables, it is possible to see how, for example, the level of education affects a housing decision. It is also possible to systematically change the independent variables in order to focus on particular groups.

American Housing Survey data is limiting in that it cannot provide detailed information about neighborhoods or other factors (i.e. schools, employment location) which affect choice. Using regional level data does little to shed light on a particular urban area. This approach may be most useful for tying the revealed preferences of particular demographic groups to housing situations in an effort to establish defensible measures of the value of particular housing programs.

3. Will this model improve planning and policy decisions by helping inform planning and policy making bodies about the housing choices made by low-income households? In particular, can the model provide information about the desirability

of targeting scarce housing resources?

The results indicate that federal housing programs increase the probability of finding affordable, adequate quality and uncrowded housing by a substantial amount. Also shown is the way federal housing programs work to improve the situation of the very low-income household. Limitations on the information available about existing programs restrict application of the methodology to the broad classifications of federal housing assistance and local housing assistance. The results suggest that further research in this area, using data that address particular housing programs, would provide a useful check on the survey-based conclusions that dominate housing literature. In view of the low incidence of any type of housing assistance (Table 4.5), whether federal or local, it is difficult not to make a strong case for targeting resources to those with the lowest incomes. Federal housing assistance increases the probability that a very low-income household finds basic housing by almost five-hundred percent while, for a household earning eighty percent of median, the probability increases by less than twenty percent. Income assistance alone may be a more effective way of helping those households approaching median income, saving the scarce housing assistance for the very low-income.

The results also suggest that, if this methodology is to be successful, careful consideration must be given to the way low-income renters perceive their housing choices and, because housing is so local market dependent, the choices available in particular areas. Stegman (1992 p.63) reinforces this point by calling for an expanded AHS survey which would include more information on the local housing stock and local markets.

Models of choice processes in the low-income renter's selection

of housing offer a better understanding to planners and policymakers with regard to relevant aspects of housing problems. This type of model can provide a basis for targeting policy decisions as well as targeting funding to areas of greatest need.

6.4 Final Comments and Suggestions for Future Research

Measuring housing needs in physical terms ignores the consequences of poor housing with respect to gaining access to other resources. There are limitations as to what a housing policy can achieve. However, a provision of the National Affordable Housing Act of 1990 is the Family Self-Sufficiency Program. This program combines housing assistance with employment training programs (Nelson and Khadduri 1992). Housing problems must be seen in the context of the entire social system which includes access to education and employment. Housing assistance alone cannot alleviate social problems. It may prove useful in the future to examine the extent to which housing deprivation coexists with other forms of deprivation such as lack of nearby jobs, poor provision of services, a lack of access to public transportation, or a neighborhood with a high crime rate.

One of the aims of this research was to integrate the effects of housing policy with measures of housing deprivation. In the future, research of this kind must include more market specific information and more specific information on the level of funding associated with individual housing assistance programs. A logical follow-up to this study would be a study which tests the sensitivity of the dependent variables. For example, the arbitrary nature of a thirty percent rent-to-income ratio ignores the problems very low-income renters have in meeting their other needs. The impact of redefining that variable on the probability of finding basic shelter (both with and without housing assistance) should be investigated. Integrating

policy variables into disaggregate models of choice will provide knowledge about the sensitivity of households to specific policy measures and aid in assessing the effectiveness of housing policy decisions.

REFERENCES

Aaron, Henry J. 1972. *Shelter and Subsidies.* Washington, D.C.: The Brookings Institution.

Anas, Alex. 1982. *Residential Location Markets and Urban Transportation.* New York: Academic Press.

Apgar, Jr., William C. 1990. Which housing policy is best? *Housing Policy Debate,* 1(1): 1-32.

Apgar, Jr., William C. 1991. *The State of the Nation's Housing 1990.* Cambridge: The Joint Center for Housing Studies at Harvard University.

Apgar, Jr., William C. 1992. Preservation of existing housing. *Housing Policy Debate,* 2(1): 187-210.

Arnott, R. 1987. Economic theory and housing. In *Handbook of Urban Economics and Regional Economics*, ed. E. S. Mills, 959-988. The Netherlands: Elsevier Science Publishers B.V..

Baer, William C. 1976. The evolution of housing indicators and housing standards. *Public Policy*, 24(3): 361-393.

Baer, William C. and Christopher B. Williamson. 1988. The fIltering of households and housing units. *Journal of Planning Literature*, 3(2): 127-152.

Ben-Akiva, Moshe and Steven R. Lerman. 1985. *Discrete Choice Analysis.* Cambridge, Massachusetts: The MIT Press.

Blank, Rebecca M. and Harvey S. Rosen. 1990. Recent trends in housing conditions among the urban poor. *Research in Urban Economics, Vol. 8.* 15-34. JAI Press Inc.

Boehm, Thomas P. 1979. *The Interrelationship Between Housing Choice and Expected Mobility.* Unpublished Ph.D. dissertation, Washington University, Saint Louis, Missouri.

Boehm, Thomas P. 1980. A Hierarchical Model of Housing Choice. *Urban Studies,* 19: 17-31.

Boehm, Thomas P. and Keith R. Ihlanfeldt. 1991. The revelation of neighborhood preferences: an n-chotomous multivariate probit approach. *Journal of Housing Economics,* 1: 33-59.

Börsch-Supan, Axel. 1990. Panel data analysis of housing choices. *Regional Science and Urban Economics,* 20: 65-82.

Börsch-Supan, Axel and John Pitkin. 1988. On discrete choice models of housing demand. *Journal of Urban Economics,* 24: 153-172.

Börsch-Supan, Axel and H. Pollakowski. 1990. Estimating housing consumption adjustments from panel data. *Journal of Urban Economics,* 27: 131-150.

Bradbury, K. A. and A. Downs, eds. 1981. *Do Housing Allowances Work?* Washington, D.C.: The Brookings Institute.

Braid, Ralph M. 1986. The comparative statics of a filtering model of housing with two income groups. *Regional Science and Urban Economics,* 16: 437-448.

Broley, C. H., M. D. Tate and L. J. Walters. 1976. Consumer preferences in housing. In *Housing in the Seventies, Working Papers 2.* Washington, D.C.: Dept. of Housing and Urban Development.

Burke, Paul. 1982. Trends in subsidized housing, 1974-1981. *HUD Report 0003034,* Washington, D.C.: Division of Housing and Demographic Analysis.

Clark, W. A. V. and J. L. Onaka. 1985. An empirical test of a joint model of residential mobility and housing choice. *Environment and Planning,* 17: 915-930.

Clemmer, Richard B. and John C. Simonson. 1983. Trends in substandard housing 1940-1980. *AREUEA Journal,* 10: 442-464.

De Leeuw, Frank and Raymond J. Struck. 1975. *The Web of Urban Housing.* Washington, D.C.: The Urban Institute.

Deurloo, M. C., F. M. Dieieman and W. V. A. Clark. 1988. Generalized log-linear models of housing choice. *Environment and Planning,* 20: 55-69.

Eekhoff, Johann. 1987. Comment. In *Reassesssing the Role of Government in the Mixed Economy,* ed. Herbert Giersch. Boulder, Colorado: Westview Press.

Ellickson, Brian. 1981. An alternative test of the hedonic theory of housing markets. *Journal of Urban Economics,* 9: 56-79.

Feins, Judith D. and Terry Saunders Lane. 1981. *How Much for Housing?* Cambridge, Massachusetts: Abt Books.

Frieden, Bernard J. and Authur P. Solomon. 1977. *The Nation's Housing: 1975 to 1985.* Cambridge, Massachusetts: Joint Center for Urban Studies of MIT and Harvard University.

Galster, George C.1987. The ecology of racial discrimination in housing. *Urban Affairs Quarterly,* 23, 84-107.

Goedert, Jeanne E. and John L. Goodman, Jr.. 1977. *Indicators of the Quality of U.S. Housing.* Washington, D.C.: The Urban Institute.

Goodman, Allen C. 1990. Demographics of individual housing demand. *Regional Science and Urban Economics,* 20: 83-102.

Hadden, Louise and Mireille Leger. 1989. *Codebook for the Annual Housing Survey Data Base.* Cambridge, Massachusetts: Abt Associates Inc.

Hammond, C. A. 1987. *The Benefits of Subsidized Housing Programs: An Intertemporal Approach.* Cambridge: Cambridge University Press.

Hanushek, Eric A. and John M. Quigley. 1982. The determinents of housing demand. *Research In Urban Economics, Vol. 2.* JAI Press, Inc.

Hays, R. Allen. 1985. *The Federal Government and Urban Housing.* Albany: State University of New York Press.

Henderson, J. V. and Y. M. Ioannides. 1987. Owner occupancy: investment vs. consuption demand. *Journal of Urban Economics*, 21: 228-241.

Hensher, David A. 1986. Sequential and full information maximum likelihood estimation of a nested logit model. *The Review of Economics and Statistics*, 68: 657-667.

Hensher, David A. and Lester W. Johnson. 1981. *Applied Discrete-Choice Modeling.* New York: John Wiley & Sons, Inc.

Jacobs, Barry G., K.R. Harney, C.L. Edson, B.S. Lane. 1982. *Guide to Federal Housing Programs.* Washington, D.C.: The Bureau of National Affairs, Inc.

Kain, John F. and William Apgar. 1977. *The Modeling of Neighborhood Change.* Cambridge: Harvard University.

Kain, John F. and John M. Quigley. 1972. Housing market discrimination, home-ownership, and savings behavior. *American Economic Review*, 62: 263-277.

Karlqvist, Anders L. 1978. In *Spatial Interaction Theory and Planning Models,* eds. A. Karlqvist L. Lundqvist, F. Snickars and J. W. Weibull. Amsterdam: North-Holland.

Kearns, A. 1990. Housing policy, deprivation and space: the case for stress areas. *Policy and Politics*, 18: 119-134.

King, M. A. 1980. An econometric model of tenure choice and demand for housing as a joint decision. *Journal of Public Economics*, 14: 137-159.

Learman, Steven R. 1979. Neighborhood choice an dtransportion services. In *The Economics of Neighborhood*, ed. D. Segal. New York: Academic Press.

Lerman, Steven R. 1983. Random utility models of spatial choice. In *Optimization and Discrete Choice in Urban Systems,* ed. B.G. Hutchinson et al. Berlin:Springer-Verlag.

Lowry, I. 1960. Filtering and housing standards: A conceptual analysis. *Land Economics,* 36: 362-370.

Luce, R. 1959. *Individual Chooice Behavior: A Theoretical Analysis.* New York: Wiley and Sons.

Maddala, C. S. 1988. *Introduction to Econometrics.* New York: Macmillan Publishing Company.

Maddala, C. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press.

Mayo, Stephen K. 1981. Theory and estimation in the economics of housing demand. *Journal of Urban Economics,* 10: 95-116.

McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics,* ed. Paul Zarembka. New Yord: Academic Press.

McFadden, Daniel. 1978. Modeling the choice of residential location. In *Spatial Interaction Theory and Planning Models,* eds. A Karlqvist, L. Lundqvist, F. Snickars and J. W. Weibull. Amsterdam: North-Holland.

McFadden, Daniel. 1981. Econometric models of probabilistic choice. In *Structural Analysis of Discrete data with Econometric Applications,* eds. C.F. Manski and D. McFadden. Cambridge, MA: MIT Press.

Mills, E. S. and A. Sullivan. 1981. Market effects. In *Do Housing Allowances Work?* eds. K. L. Bradbury and A. Downs. Washington, D.C.: The Brookings Institute.

Murray, M. P. 1980. Tenant benefits in alternative federal housing programs. *Urban Studies,* 17: 25-34.

Muth, R. F. 1975. *Urban Economic Problems.* New York: Harper and Row.

Nelson, Kathryn P. 1992. Data for comprehensive housing affordability strategies. *Journal of the American Planning Association,* 58: 85-102.

Nelson, Kathryn P. and Jill Khadduri. 1992. To whom shoould limited housing resources be directed? *Housing Policy Debate*, 3: 1-35.

Newman, Sandra and Ann B. Schnare. 1988. *Subsidizing Shelter.* Washington, D.C.: The Urban Institute.

Newman, Sandra J. and Raymond J. Struyk. 1983. Housing and poverty. *The Review of Econoomics and Statistics,* 65: 243-253.

Olsen, E. O. 1982. Housing programs and the forgotten taxpayer. *Public Interest*, 66: 97-109.

Onaka, J. and W. A. V. Clark. 1983. A disaggregate model of residential mobility and housing choice. *Geographical Analysis*, 15: 287-304.

Polinski, A. M. 1977. The demand for housing: a study in specification and grouping. *Econometrica,* 45: 447-461.

Quigley, John M. 1976. Housing demand in the short run: an analysis of polytomous choice. *Explorations in Economic Research*, 3: 76-102

Quigley, John M. 1983. Estimates of a more general model of consumer choice in the housing markets. In *The Urban Economy in Housing* ed. R.E. Grierson. Lexington, MS: Lexington Books.

Qugley, John M. 1985. Consumer choice of dwelling, neighborhood, and public service. *Regional Science and Urban Economics*, 15: 41-63.

Ratcliff, R. 1949. *Urban Land Economics.* New York: McGraw-Hill.

Roistacher, Elizabeth A. 1981. Introduction. In *Rental Housing: Is There A Crisis?* eds. J.C. Weicher, K.E. Villani and E.A. Roistacher, 1-8. Washington, D.C.: The Urban Institute Press.

Rosen, Harvey S. 1985. Housing subsidies: effects on housing decisions, efficiency, and equity. In *Handbook of Public Economics, vol I.*, eds. A.J. Auerbach and M. Feldstein, 375-420. Amsterdam: Elsvier Science Publishers B.V.

Rouwendal, Jan. 1989. *Choice and Allocation Models for the Housing Market.* Dordrecht: Kluwer Academic Press.

Sa-Aadu, Jarjisu. 1984. Another look at the economics of demand-side versus supply-side strategies in low-income housing. *AREUEA Journal*, 12(4): 427-460.

Salins, Peter D. 1987. Can we ensure that all Americans are well housed? In *Housing America's Poor*, ed. Peter D. Salins, 175-194. Chapel Hill: The University of North Carolina Press.

Spain, Daphne. 1990. The Effect of Residential Mobility and Household Composition on Housing Quality. *Urban Affairs Quarterly*, 25(4): 659-683.

Stegman, Michael A. 1992. Comment. *Housing Policy Debate,* 3: 7-66.

Stone, Michael E. 1990. *One-Third of a Nation, A New Look at Housing Affordability in America.* Washington, D.C.: Economic Policy Institute.

Stiglitz, Joseph E. 1988. *Economics of the Public Sector.* New York: W. W. Norton & Company.

Swartz, David C., R. C. Ferlaunto and D. N. Hoffman. 1988. *A New Housing Policy for America.* Philadelphia: Temple University Press.

Sweeney, James L. 1974. A commodity hierarchy model of the rental housing market. *Journal of Urban Economics,* 1: 288-323.

Thurstone, L. 1927. A law of comparative judgement. *Psychological Review,* 34: 273-286.

Train, Kenneth. 1986. *Qualitative Choice Analysis: Theory, Economics and An Application to Automobile Demand.* Cambridge: MIT Press.

Urban Land Institute. 1992. *Land Use Digest,* 25:4.

Van Lierop, Wal. 1986. *Spatial Interaction Modeling and Residential Choice Analysis.* Aldershot, England: Goower Publishing Company Ltd.

Van Lierop, Wal and A. Rima. 1985. Trends in prospects for qualitative disaggregate spatial choice models. In *Measuring the Unmeasurable* eds. Peter H. Nijkamp, Leitmer and N. Wrigley. Dordrecht: Martinus Nijhoff Publishers.

Venti, Stever F. and David A. Wise. 1984. Moving and housing expenditure: transaction costs and disequilibrium. *Journal of Public Economics,* 23: 207-239.

Vitaliano, Donald F. 1983. Public housing and slums: cure or cause? *Urban Studies,* 20: 173-163.

Weicher, John. 1979. Urban housing policy. In *Current Issues in Urban Economics* eds. P. Mieszkowski and M. Straszheim. Baltimore: The John Hopkins University Press.

Weicher, John C. 1987. Private production: has the rising tide lifted all boats? In *Housing America's Poor* ed. Peter D. Salins. Chapel Hill and London: The University of North Carolina Press.

Weicher,John C. and Thomas G. Thibodeau. 1988. Filtering and housing markets: an empirical analysis. *Journal of Urban Economics,* 23: 21-40.

Wieand, Kenneth F. 1983. The preformance of annual housing survey quality mesures in explaining dwellling rentals in 20 metropolitian areas. *AREUEA Journal,* 11(1): 45-68.

Wrigley, Neil. 1985. *Categorical Data Analysis For Geographers and Environmental Scientists.* New York: Longman Group Limited.

Wrigley, N. and P. Longley. 1984. Discrete choice modeling in urban analysis. In *Geography and the Urban Environment, Vol. VI* 45-94, eds. D. T. Herbert and R. J. Johnston. London: John Wiley.

APPENDIX A
DEFINITION OF HOUSING QUALITY

DEFINITION OF HOUSING QUALITY

Unit lacks or shares complete plumbing facilities.

Unit lacks adequate provision for sewage disposal. The unit must be connected with a public sewer, septic tank, cesspool, or chemical toilet.

Unit lacks or shares complete kitchen facilities.

Has two or more structural problems:
- Leaking roof.
- Leaking basement.
- Open cracks or holes in interior walls or ceilings.
- Holes in the interior floors.
- Peeling paint or broken plaster over one square foot on an interior wall.
- Evidence of mice or rats in the last 90 days.

Has two or more common area problems:
- No working light fixture in common hallway.
- Loose, broken, or missing stairs.
- Broken or missing stair railings.
- No elevator in building (for units two or more floors from main building entrance in building four or more stories high).

Unit is heated mainly by unvented room heaters which burn gas, oil, or kerosene.

Unit has had three or more toilet breakdowns or six hours or more in the past 90 days.

Unit had three or more heating breakdowns lasting six hours or more last winter.

Lacks electricity.

One or more rooms with out a working wall outlet.

Fuses blown or circuit breakers tripped three or more times during last 90 days.

Exposed wiring in house.

Source: Definition used in Newman and Schnare 1988, p. 105.

VITA

Margaret S. Murray was born in Iowa Falls, Iowa on July 19, 1940. She received a Bachelor of Arts in Education from the University of Northern Iowa in 1962 and a Master of Science in Education from the University of Kansas in 1970. In 1979, she received a Master of Business Administration degree from the University of Oklahoma. Ms. Murray's work experience includes real estate sales and university teaching.

Degree requirements for the Ph.D. in Environmental Design and Planning with an emphasis in housing policy were completed in November of 1992. She was the recipient of a Presidential fellowship at VPI & SU. Ms. Murray is a member of the American Planning Association, a student member of the American Real Estate and Urban Economics Association and an associate member of the Urban Land Institute.

Margaret S. Murray