Another specification issue that warrants discussion is the stratification of data samples. After specifying a general model of tenure choice, several studies estimate the model separately for subsets of their samples. Because the objective of this study is to investigate the systematic determinants of the differential probability of homeownership by racial or ethnic groups, we base our analyses on the assumption that the impacts of other (i.e., nonracial and nonethnic) explanatory variables on the decision to own or rent may be different for various groups and thus stratify our sample.

Many analysts view choice of household tenure as a decision that can be considered independently of other choices within a broader context. However, theoretical analyses more often argue that tenure can best be understood as an element of broader choices, including mobility, housing consumption, and labor supply changes (see Haurin 1991). We view tenure choice as an element of both consumption and investment decision making. We follow Goodman's (1988) innovation of separating the consumption and investment motives in the tenure choice decision through the formulation of owner-renter price ratios and value-rent ratios.

All empirical studies of the choice of tenure have applied multivariate techniques to test the usefulness of various independent variables in explaining a household's tenure choice. They differ, however, in the structure and type of the multivariate model estimated.[7] Most employ logit as their estimating process.

An additional model specification issue is determining which set of specific independent variables to include and how to define them for operational purposes. Turner and O'Neal (l986) discuss a wide range of variables that analysts have investigated. These fall into five categories: (1) income and wealth, (2) life-cycle status, (3) race and ethnicity, (4) price and other market factors, and (5) location and neighborhood attributes. Location and neighborhood attributes do not have to enter the tenure choice equation directly if price indices are measured to control for them.

Income and wealth. Household income and wealth play a prominent role in determining a household's ability to own a home. Virtually all studies of tenure choice have found income a statistically significant

---

[7] For example, Kain and Quigley (1975) use generalized least squares to provide more efficient estimates; Roistacher and Goodman (1976) test their results by using logistic formulation; Krumm (1984) uses logit analysis to estimate the probability that non-white and white householders would continually rent; and Henderson and Ioannides (1987) use probit to analyze the household tenure decision.

determinant of the probability of ownership. The most analytically desirable measure of income is household permanent income because it reflects longer-term income capacity, including that from nonhuman capital. Permanent income, however, is not directly observed. Because a long time series on income for each household is not available, we adopted a procedure used by Follain (1979) and Goodman and Kawai (1986) to construct a measure of permanent income.

The literature also emphasizes the two routes through which changes in current wealth affect tenure choice. First is the standard "wealth effect," which in theory is equivalent to a change in discounted lifetime earnings (Haurin 1991; Zorn 1988). Second is mortgage lenders' requirement that current wealth must be sufficient for a down payment on the desired house. The lack of current wealth increases the probability that a household will elect to rent. Linneman and Wachter (1989) identify the critical importance of the down payment on tenure choice. They note that controlling for current wealth when analyzing tenure choice empirically confuses the down payment constraint and wealth effects. In our model, we do not control for down payment constraint effects because AHS does not include sufficiently detailed measures of wealth. However, we do include estimates of wealth in our permanent income estimation. Thus, our measure of permanent income includes wealth and wage income.

Life-cycle status. The second group of variables encompasses measures of life-cycle status, including age, family type, and household size.[8] The use of sociodeomographic variables in housing demand analysis is extensive and varied, as evidenced by Mayo's (1981) instructive review. Although recent work on the role of sociodemographic factors in tenure choice remains incomplete (Barnes and Gillingham 1984; Pollak and Wales 1981), analysts agree that an understanding of permanent income provides insight into many of the sociodemographic effects. Goodman and Kawai (1984) show how the impacts of education, age, and training on permanent income are likely to be nonmonotonic.[9] More recently,

---

[8] There are numerous theories about why life-cycle status plays a role in the tenure choice decision. One hypothesis is that different types of households exhibit different demand for housing relative to other goods and that homeownership may be associated with higher levels of housing consumption. Another hypothesis posits that certain types of housing—such as single-family suburban dwellings—are more likely to be available for sale rather than rent and that groups of households preferring these types of dwellings will therefore be more likely to own. A third hypothesis holds that expectations about the immediate future with respect to moving costs and the possible investment characteristics of the housing bundle may affect tenure choice.

[9] If life-cycle effects are ignored in the permanent income formulation, then nonmonotonic effects would tend to persist in the tenure choice regressions through the coefficients on permanent income.

Goodman (1990) provides a systematic investigation of the role of demographic factors in housing demand specification. We control for nonmonotonic effects of sociodemographic effects on permanent income.

Housing price indices. The final set of market variables includes the relative price of owning compared with renting as one of the traditional determinants of the likelihood of ownership. The relative price data entered as explanatory variables should refer to a constant quantity of housing services (or level of housing quality) expressed over time, across locations, in terms of consumption and investment attributes, and as an ownership versus rental option. It is not easy to cast the quantity of housing services along such dimensions. The approaches that analysts have used to measure relative housing price include the concepts of observed rents, rent index (Shelton 1978), and user cost (Hendershott and Shilling 1982).

Analysis of tenure choice focuses on the individual household engaged in a decision related to individual houses. Hence, it is important to devise house-specific measures that summarize dwelling unit investment possibilities while capturing the dwelling's consumption components and constant quality. Goodman (1988) introduced into the literature two ratios: a house-specific ratio and a market-specific ratio. The ratio of house value to renter value is the house-specific measure that indicates the relationship between consumption and investment characteristics and between renting and owning. With high expected housing price appreciation, which encourages owner-occupancy, the variable will take on high values.[10] Thus this variable is expected to be positively related to homeownership. The ratio of owner price to renter price is the market-specific measure that controls for the quality of houses across markets. This ratio is expected to have an inverse relationship to owner occupancy.

Estimating the models

The impact of racial (ethnic) variables on tenure choice and housing demand has two aspects. The first may be called tastes, but includes some choices that are based not on preferences but on discrimination.

[10] The value-rent ratio relates asset value to the rent that would have been charged. A high value-rent ratio suggests a market-indicated expectation of a capital gain. It is thus a component of the user cost of housing. Goodman (1988) provides a more complete presentation of the derivation and analysis of the model used here.

Blacks or Hispanics may, for example, have tastes for less housing, or they may be excluded or restricted from some segment of the housing market by discriminatory practices.[11] Failure to account for these differences implies that all households, irrespective of race or ethnicity, have the same utility (and hence the same demand or same propensity for ownership) functions or that all households receive identical treatment in the housing market. Such may theoretically be the case, but it is a highly restrictive assumption.

The second aspect involves the treatment of racial (ethnic) variables in the tenure choice process. If tenure choice is determined by income, the relative cost of owning versus renting, and demographic factors, then it is reasonable to believe that the impacts of changes in these factors may vary with respect to racial or ethnic status.

Analysts have proposed several forms for the inclusion of racial (ethnic) variables in tenure choice equations. Consider a simple tenure choice indicator, which involves only vectors of prices and income and demographic factors and excludes race or ethnic variables. Translation involves the addition of dummy variables to indicate racial or ethnic status that does not interact with the prices, incomes, and demographic terms of the tenure choice equation. Thus, parameter estimates are invariant to the racial (ethnic terms), although elasticities can vary.

Scaling redefines the parameters of the tenure choice as functions of racial (ethnic) variables. The two approaches can be combined by first scaling and then translating (often referred to as a Gorman specification) or by translating and then scaling [defined by Pollak and Wales (1981), as a reverse Gorman specification]. In the logistic formulation of tenure choice equations estimated in this study, these forms are nested and permit rigorous statistical testing.

Consider the formulation of a tenure choice indicator I:

$$I = \alpha Y + \beta P + \psi D + \theta_1 R + \theta_2 RY + \theta_3 RP + \theta_4 RD + \xi \qquad (1)$$

where Y is the appropriate measure of income vector, P is the relevant price vector, D is the vector of demographic terms, and R is the vector of racial (ethnic) terms. In this system, $\theta_1 = \theta_2 = \theta_3 = \theta_4 = 0$. Linear translation restricts parameters θ_2, θ_3 and θ_4 to 0. Linear scaling restricts θ_1, to while Gorman's method (which involves

[11] There is considerable direct evidence on the discriminatory treatment of blacks (Turner et al. 1991), while there is little or no ethnographic evidence to support the view that blacks inherently have tastes for less housing, even if they consume less housing.

scaling and then translating) allows all four parameters to vary. The various formulations in Equation 1 are nested within the Gorman (henceforth, extended) formulations, thus allowing testing for the inclusion of variables in the logit by using likelihood ratio tests. Equation 1 provides a useful format for classifying the treatment of racial and ethnic effects in other tenure choice studies. Race, that is, being black, is usually negatively related to homeownership. Most studies use translation methods ($\theta_2 = \theta_3 = \theta_4 = 0$), although some use scaling ($\theta_1 = 0$). None of the studies systematically examines and compares the various formulations of the racial impacts as sketched out in this study.

Following a method formalized by Goodman (1988), tenure choice can be postulated to be a function of income (Y) and demographic (D) variables, the relative price ratio of owner and renter housing (P_o/P_r), and the value-rent ratio (V as a measure of asset viability). Thus,

$$\Gamma = \Gamma(Y, P_o/P_r, V, D). \tag{2}$$

In estimating Equation 2 for the various subsamples analyzed in this study, we apply the specification of Equation 1 to logit Equations 3 through 7. To allow for the hierarchy of tests required to isolate direct and indirect race (ethnic) effects of racial discrimination, we estimate five logit equations as follows:

$$\Gamma^{w}_{1} = \Gamma(Y, P_o/P_r, V, D) \text{ (white-only sample)}, \tag{3}$$

$$\Gamma^{b}_{2} = \Gamma(Y, P_o/P_r, V, D) \text{ (black-only sample)}, \tag{4}$$

$$\Gamma_3 = \Gamma(Y, P_o/P_r, V, D) \text{ (pooled, without race dummy)}, \tag{5}$$

$$\Gamma_4 = \Gamma(Y, P_o/P_r, V, D, R) \text{ (pooled, with race dummy), and} \tag{6}$$

$$\Gamma_5 = \Gamma(Y, P_o/P_r, V, D, R, RY, R[P_o/P_r], RV, RD) \tag{7}$$

(pooled, with full interaction terms with race).

We repeat this process for Hispanic and non-Hispanic subsamples.

Decomposition of group mean differences

The model we employ decomposes mean differences between two racial (ethnic) groups into endowment and residual differences. The endowment effects result from household characteristics (e.g., age of head, family size, marital status, education, income), price, and other market factors faced by individual households. The residual difference may result from housing market racial (ethnic) discrimination. We are aware of three applications of the decomposition technique to home-tenure-choice studies (Ladenson 1978; Silberman, Yochum, and Ihlanfeldt 1982; Weinberg 1978).

Using the decomposition technique to account for the interaction effects in home-tenure-choice equations for racial groups is an adaptation of labor market studies.[12] Despite its widespread use, only recently have researchers explicitly considered the stochastic nature of the components into which group mean differences are decomposed (Jackson and Lindley 1989). Hitherto, decomposing group mean differences into endowment and residual differences treated estimated coefficients as deterministic. Indeed, most of the results are presented as simple percentage differences with imprecise references to the magnitude of these differences. However, these estimates are affected by sampling variability that is, in part, stochastic. Accordingly, Jackson and Lindley (1989) developed an empirical Chow-type test for evaluating the statistical significance of these differences.

This joint test of the two components of the residual difference—the constant effect and the coefficient effect—allows the statistical consideration of each component of the residual difference as a potential source of discrimination. The variation of the decomposition model employed in this paper is the adaptation of the Blinder technique as applied to the problem of absenteeism (Ault et al. 1991). It explicitly accounts for idiosyncratic data with a dichotomous dependent variable by employing tobit analysis and by decomposing differences in the expected mean rather than differences in the observed sample mean. Even though the adaptation of the Blinder technique to the problem of absenteeism was developed by using the tobit model, the statistical testing procedures are directly applicable to logit analysis.

White-black mean probability differences

The decomposition technique is discussed in terms of using the comparison between white and black households. The behavioral

[12] Blinder (1973) and Oaxaca (1973) independently introduced the technique into the labor market literature. Since then, the methodology has been widely used in labor market studies to measure the extent of wage discrimination (Borjas 1978; Butler 1982; Duncan and Hoffman 1983; Filer 1983; Flannagan 1974; Haig 1982; Hirsh and Leppel 1982; Jackson and Lindley 1989; Link and Ratledge 1975; Malkiel and Malkiel 1973; Mincer and Polachek 1978). This technique has also been used in locational-site choice studies (Dunlevy 1983). In fact, the decomposition technique is not limited to economics; it was applied in the sociology literature even before it was used in economics (Althauser and Wigler 1972).

models for tenure choice are stated as Equations 3 through 7. The mean probability of homeownership for each group ($\acute{\Gamma}^w$ and $\acute{\Gamma}^b$ for white and black households, respectively) is given by Equation 8 or by the sum of the product of the group's vector of coefficients (R) with the means of the vector of their respective explanatory variables (ω). For a sample composed only of blacks:

$$\acute{\Gamma}^b = \gamma^b + \Omega^b \omega^b. \tag{8}$$

The b superscript denotes that the results refer to the black sample. For a sample composed only of whites, the superscript is changed to w. Viewing mean probability of homeownership in this way allows the use of the decomposition technique.

Specifically, the technique allows us to partition the difference in average probability of homeownership ($\acute{\Gamma}^w - \acute{\Gamma}^b$) into a portion attributable to differences in average "endowments" of tenure choice determinants and a portion that may be attributable to the direct and indirect effects of racial discrimination per se.

The procedure constructs a hypothetical mean probability of homeownership ($\acute{\Gamma}^h$) for the average black household if its propensity for homeownership responded to changes in its determinants in a manner and magnitude identical to that of a white household. The procedure is accomplished by multiplying the white household's coefficient values by the black household's mean characteristic values and summing:

$$\acute{\Gamma}^h = \gamma^w + \Omega^w \omega^b. \tag{9}$$

The partition amounts to adding and subtracting the hypothetical mean from the differential of the mean probability of homeownership for white and black households, such that

$$(\acute{\Gamma}^w - \acute{\Gamma}^b) = (\acute{\Gamma}^w - \acute{\Gamma}^h) + (\acute{\Gamma}^h - \acute{\Gamma}^b). \tag{10}$$

The total difference in mean probability of homeownership ($\acute{\Gamma}^w - \acute{\Gamma}^b$) given by Equation 10 can also be expressed as the sum of the endowment effect: the constant effect, and the coefficient effect, respectively, as given in Equation 11:

$$(\acute{\Gamma}^w - \acute{\Gamma}^b) = [\Omega^w(\omega^w - \omega^b)] + (\gamma^w - \gamma^b) + [\omega^b(\Omega^w - \Omega^b)]. \tag{11}$$

The first term on the right in Equation 10 is called the "endowment effect." The intuition for this identification can be seen from Equation 11, where the term is rewritten as the weighted sum of the mean differences in corresponding determinants of the probability of homeownership (where the weights are the white response coefficients). As such, this portion of the mean probability differential is the result of factors that determine the differences in mean characteristics of white and black households. Similarly, the second term on the right in Equation 10 is the residual difference. This difference between the mean probability of homeownership of white households and our hypothetical black household ($\acute{\Gamma}^{w} - \acute{\Gamma}^{h}$) is a sum of two components as seen in Equation 11: the constant and coefficient effects. The residual difference may be attributable to nonmeasured factors as well as racial discrimination in housing and mortage markets per se.

Statistical testing procedures

The fundamentals of the statistical testing involved in adapting the decomposition methodology as employed in this paper are contained in Ault et al. (1991). The sum of the log-likelihood functions for the white-only and black-only logit Equations 3 and 4 can be viewed as the unconstrained log-likelihood function for the sample as both sets of logit coefficients and variances are allowed to differ between the two groups; thus, $L_u = L^w + L^b$. The log-likelihood function from Equation 5 can be viewed as that of a fully constrained model (L_f); the pooling procedure constrains the logit coefficients and variances to be the same for the two groups.

The logit Equations 6 and 7 allow us to partition the residual difference into its component effects and to test for equality of variances between the two groups. Equation 6 employs the pooled sample with a group dummy for race. This model constrains the slope coefficients and variances to equality but allows the logit constant terms to differ for white and black households; the log-likelihood of this model is termed L_c. Equation 7 allows slope coefficients as well as the intercept to differ between groups as only the variances are constrained to equality. Interaction terms are added for each explanatory variable (created as the product of the group dummy with each of the individual characteristics explaining tenure choice probability) to the above dummy variable model; the log-likelihood function of this model is termed L_v. Finally (assuming equal variances), it is important to note that the logit regression coefficients on the interaction variables in this model are the differences in the logit regression coefficients on corresponding variables in the white-only and black-only logit.

Following Ault et al. (1991), the results of estimating these five logit models allow us to conduct a set of likelihood ratio tests and asymptotic t-tests that shed considerable light on the effect of race on homeownership. These tests are not valid unless the variances of the underlying models are equal. Furthermore, pooling data (as required by three of our models) in the presence of unequal variances is not legitimate because it induces heteroscedasticity. A likelihood ratio test of the null hypothesis $H_o : \sigma_w = \sigma_b$ can be conducted by using the statistic $X_1 = -2(L\text{v} - L\text{u})$, which is distributed chi square with 1 degree of freedom. Rejection of the null hypothesis requires a correction to ensure homoscedasticity before conducting the remaining tests. A test of a sufficient condition for the significance of the residual difference is a test of equality of (normalized) logit coefficients between white and black households. This test can be conducted by applying a likelihood ratio procedure that uses the statistic $X_2 = -2(L_f - L_u)$, which is distributed chi square with k+1 degrees of freedom, where k is the number of explanatory variables, excluding the constant. In our basic model, rejection of the null hypothesis implies that the combined direct and indirect effects of racial discrimination significantly affect homeownership. The obvious question in this case is whether the difference is due to the constant effect or the coefficient effect. The joint significance of the slope coefficients of the coefficient effect can be tested by using the likelihood ratio statistic $X_3 = 2(L_c - L_u)$, which is distributed chi square with k degrees of freedom.

The most obvious test of the significance of the constant effect is an asymptotic t-test of the significance of the coefficient on the race dummy in the fully interactive model. Even if the components of the residual difference are jointly insignificant, examination of the individual significance of the various coefficient differences (interaction term coefficients) can indicate which linkages are important in transmitting any extant differential effects of these variables on homeownership for white and black households. Within this hierarchical scheme of tests, (joint) insignificance of the residual difference coupled with individual insignificance of both the race dummy and all interaction variables in the fully interactive model would provide strong evidence that discrimination in the racial housing market per se does not affect homeownership. The identical testing procedure is repeated for the difference in the mean probability of ownership between Hispanic and non-Hispanic households.

Empirical results

Sample summary statistics

The data used in this paper are taken from the 1989 National Sample of the AHS that was drawn from the 1980 census. The AHS collects data every other year on the nation's people and their homes (between 50,000 and 80,000 homes); it draws a single sample that grows slightly larger from survey to survey to include new homes. The national interviews are conducted in the fall. The sample is composed of 87 percent white households and 10 percent black; the remaining 3 percent is largely Asian households (68 percent). In terms of ethnic composition, Hispanic households account for 5.7 percent of all households; of the Hispanic households, 2.7 percent are black. Therefore, most Hispanic and non-Hispanic households are headed by whites. The means and standard deviations for the variables used in the regression analysis are presented in table 2. Using AHS data, we construct two sets of economic variables as follows:

Table 2. **Sample Statistics for Racial and Ethnic Subgroups**

	Means (Standard Deviations)					
Variable	White	Black	t-value of mean diff.	Hispanic	Non-Hispanic	t-value of mean diff.
HINCOME	32,419	23,450	21.40	26,509	31,689	9.14
	(28,236)	(21,157)		(23,394)	(26,012)	
PINCOME	30,329	21,540	33.82	24,478	29,617	16.35
	(15,422)	(13,440)		(12,812)	(15,652)	
TINCOME	2,090	1,910	0.61	2,031	2,072	0.10
	(18,819)	(15,033)		(16,773)	(18,574)	
PRRATIO	185.17	184.85	0.40	205.17	184.53	-19.10
	(36.70)	(42.63)		(43.42)	(37.20)	
VALRENT	112.44	107.09	5.86	121.75	111.44	-8.37
	(46.94)	(48.31)		(51.34)	(47.08)	
AGE	47.01	45.86	3.62	41.23	47.08	15.55
	(17.99)	(16.55)		(15.42)	(17.84)	
PER	2.53	2.78	-8.16	3.31	2.53	-18.21
	(1.42)	(1.64)		(1.80)	(1.43)	
MARRIED	0.55	0.37	19.67	0.56	0.52	-3.31
	(0.50)	(0.48)		(0.50)	(0.50)	
SEX	0.68	0.49	20.19	0.65	0.66	0.87
	(0.47)	(0.50)		(0.48)	(0.47)	

Note: Variable definitions for tenure choice logit regression
TENURE: 1 if owner, otherwise 0
HINCOME: household income
PINCOME: permanent income
TINCOME: transitory income
PRRATIO: price ratio (price of owner/price of renter)
VALRENT: value-rent ratio (predicted value/predicted gross rent)
AGE: age of householder
SEX: 1 if male householder, otherwise 0
PER: number of persons in the household
MARRIED: 1 if married householder, otherwise 0
BLACK: 1 if black head of household, otherwise 0
SPAN: 1 if Hispanic head of household, otherwise 0

Interaction Variables

BLA stands for BLACK
SPA stands for SPAN
PIN stands for PINCOME
TIN stands for TINCOME
PRR stands for PRRATIO
VAL stands for VALRENT
AGE stands for AGE
MAR stands for MARRIED
SEX stands for SEX