CHAPTER III: ESTIMATION RESULTS

This chapter presents the results of estimating the model described in chapter II to the fifty-nine SMSA Annual Housing Surveys currently available. The actual estimates are available in a separate data appendix (available from the authors). Exhibit 4 presents a typical set of results, for Washington, D.C.

It is difficult to report over five thousand regression coefficients without sacrificing some verve and clarity of exposition. For this reason, we attempt to summarize these results in the following manner. First, we discuss the overall performance of the hedonic model, focusing mainly on its predictive power, and whether individual coefficient estimates are consistent with our prior beliefs. Second, we address the variation of the estimates among SMSAs. It is desirable that the estimates exhibit some stability, yet if the differences in estimates can all be explained away as error, or statistical "noise," there is no point in estimating separate indexes for different markets. A central premise of this work is that housing markets are local and diverse. Estimates of the market clearing hedonic prices should vary because of this diversity.

SECTION 3.1: DOES THE BASIC EQUATION MAKE SENSE?

The answer to this question is based upon three criteria: (1) the explanatory power of the estimated equations; (2) the statistical significance of the coefficient estimates and whether their signs (positive or negative) conform to a priori beliefs, and (3) an examination of the residuals. Unless the estimates of the basic equation explain variations in rents and values for most SMSAs, the estimates are of little use. Likewise, if few of the coefficient estimates are statistically significant, it is unlikely that an analysis of them individually would produce much insight into the operations of housing markets. These criteria are necessary (though not sufficient) to guarantee the validity of the basic equation. This chapter examines the first two criteria; residual analysis is deferred until Chapter IV.

Explanatory Power

Explanatory power is measured by three related statistics: R^2, the F statistic and the standard error of the equation. All give insight into the predictive power of the full equation.

The hedonic model succeeds in accounting for much of the observed variation in rent and value (Exhibit 5). The median multiple correlation coefficient(R^2) is .61 for the 59 owner equations and .67 for the renter equations. The Honolulu owner regression performs poorly relative to other regresions, with an R^2 of only .32. Other than this outlier, the R^2 for owners ranges from .49 for Providence to .74 for Memphis. For renters, R^2 ranges from .52 for Newark to .82 for Raleigh.

Exhibit 6 presents stem-and-leaf plots of the R^2 statistic, by tenure type. A stem-and-leaf plot is a convenient way to summarize the distribution of any set of numbers.[1] To read the plot, find the first digit of the number (here, R^2) on the vertical axis (called the stem). The second digit is read off the leaf which is to the right of the stem, row by row. For example, the bottom row of the renter plot:

.5 22223

tells us that there are four regressions with R^2 of .52, and one with an R^2 of .53.

Forty-one owner regressions and 47 renter regressions have R^2 greater than or equal to .6.

Although all but two R^2s are greater than .50 and most are between .60 and .70, it is interesting to examine their distribution. Two patterns emerge from Exhibits 5 and 6. First, there is a positive correlation between the R^2 for each tenure type, city by city (.66). Second, in 41 SMSAs renter R^2 is greater than owner R^2. The amount of variation explained does depend somewhat on the tenure group, but the differences in R^2 are not overwhelming.

The standard errors of the estimates of the natural log of value and rent are presented in Exhibit 5. The standard error is a measure of the hedonic regression's ability to predict value and rent. Since the equation predicts either the log of value or the log of rent, standard errors are also in logarithms. They can be interpreted as

---

1. See John W. Tukey (1977) for more on the uses of these plots.

an average percentage error in predicted rent or value, with larger errors receiving larger weights because the deviations are squared before averaging.[1]

The standard errors of the owner regressions range from 20 percent (Paterson) to 37 percent (Birmingham). Most (36) are less than 30 percent; and the average is about 29 percent. The dollar value of the average standard error, evaluated at the average value, is roughly nine thousand dollars.

The standard errors of the renter regressions are distributed similarly to the owner estimates, although the renter estimates are more tightly grouped below 30 percent. The range is from 19 percent (Las Vegas) to 35 percent (Honolulu). The average is about 25 perce cent. The dollar value of the average standard error is, roughly, $40.

Finally, an F-test shows that all 118 regressions are statistically significant. That is, we reject the hypothesis that the observed multiple correlation is due to chance. Both renter and owner equations perform well by this criterion.

Statistical Significance of the Individual Coefficient Estimates

In this section, the statistical significance of individual coefficient estimates is examined. Also, the signs and magnitudes of the average coefficient estimates are examined. Exhibit 7 contains summary information about the individual estimates for renters, includ-

---

1. The interpretation of the standard error as a percentage is an approximation, because of the logarithmic transformation. That is, the (weighted) average of the logarithmic error term is not the same as the logarithm of the (weighted) average dollar error in predicted rents or values.

ing the mean and median estimates of each coefficient, the range of the estimates and the number of times the estimates are significant.[1] Exhibit 8 contains the same information for owners.

Some general observations should be made. First, most of the coefficient estimates are significant at any commonly used level of significance. For example, for renters, 76 percent of the estimates have t-statistics greater than 1; 63 percent have t′s greater than 1.64 (significant level = .1 for 2-tailed test); and 48 percent have t′s greater than 2.58 (significant level = .01).

Second, our method of using indicator variables and extremely flexible functional forms (extra quadratic and cubic terms) means that many interesting hypotheses regarding coefficients should be tested using an F-test, not a t-test. An F-test allows us to test

---

[1] By statistically significant, we mean estimates which have a t-statistic greater than 1.64 or less than -1.64. This corresponds to a significance level of approximately .1 for a 2-tailed test. A significance level of .1 means that, over many trials, we could expect to reject the hypothesis that the coefficient is zero, when it is in fact is zero, one time out of ten. Our choice of significance level is necessarily somewhat arbitrary. In many econometric contexts, confidence intervals convey more useful information than simply reporting rejection of a null hypothesis. We choose to report the number of times coefficients are significant at an arbitrary level (.1) because it is difficult to summarize 59 confidence intervals for any single coefficient.

These and other hypothesis tests used in this paper assume that the errors in the models are normally distributed. In fact, analysis of the estimated error terms (residuals) show that the error term is probably not normal, but is slightly skewed. Several studies have shown that t-tests are not terribly sensitive to moderate deviations from normality, as long as the error distribution is still bell-shaped (see Theil 1971, pp. 615-16). This means our significance levels are not exact but reasonable approximations.

A second reason hypothesis tests and confidence intervals constructed from these results are not exact is that we estimate a model we have constructed using a specification search (see Leamer, 1978). This means that the probability of an incorrect decision (rejecting or not rejecting a hypothesis) is actually somewhat higher than classical statistical methods indicate, because the data have helped form the model.

the joint effect of several variables. For example, all four variables measuring the age of the structure could yield low t-statistics, while the effect of age on value remains strong.[1] Several F-tests of joint hypotheses have been included and will be discussed below.

The average estimates of the coefficients are almost always consistent with a priori considerations. For example, the average estimates of the coefficients of the number of baths, the number of rooms, and the number of bedrooms are positive. The average estimates of the coefficients of the house age, the occupant's rating of the street, and the indicator variable measuring the presence of deteriorated housing are all negative.

Of course, many variables occasionally exhibit coefficients that are not of the expected sign but are "significant," or, more properly, whose t-statistic has an absolute value greater than the chosen cutoff ($|t| > 1.64$). Some anomalies are expected, in a statistical sense, since the coefficient estimates are random variables, and we estimate several thousand.

Several variables have the unexpected sign more often than should be the result of chance. For example, for renters the coefficients of structure-type variables such as SFATT and SFDET (single-family attached and detached) have signs which indicate that these variables are probably picking up locational effects rather than structural effects. That is, we expect that if otherwise identical

---

[1] That is, large values of AGE1, AGE1SQ, AGE1CB, and DAGE occur together. The t-statistics measure the effect of each variable, which may be diminished by including several collinear age variables, while their joint effect is strong. The F-test measures this joint effect.

single-family houses and multi-family units were located side by side, the single-family dwellings would command higher rents. The consistently negative coefficients for SFATT and SFDET, and the positive coefficient for NGT50 (greater than 50 units) indicate that these variables are probably proxies for the price of location, since single-family dwellings are concentrated where land prices are lowest.

Only a few variables other than structure-type variables consistently exhibit puzzling signs. NORAD, rooms without heating equipment, is negative and significant in 17 renter and 34 owner regressions, but is the wrong sign and significant in 6 renter regressions and 11 owner regressions. Now, the implicit prices of heating equipment might vary with climate, but the SMSAs where people pay more for dwellings with unheated rooms include northern cities such as Albany and Boston as well as Los Angeles and Miami. POORN, poor neighborhood rating, has the wrong sign and a t-statistic greater than 1.64 in 6 owner regressions, but conforms to expectations in 18 SMSAs.

SECTION 3.2: EXAMINING THE STRUCTURAL COEFFICIENTS

Examining Individual Coefficient Estimates

In the next four sections we discuss the distributions of estimated coefficients for the renter and owner models. The strategy followed is to identify the average estimated coefficient, the spread of the estimated coefficients about the average, and important outliers. The SMSAs with high or low estimated coefficients are examined for obvious similarities in location or growth rate. We indicate the number of statistically significant estimates as well as those that are inconsistent with our prior beliefs. The discussion of the distributions of estimated coefficients basically follows the list of regressors as they appear in Exhibit 2.

The first category includes all coefficients relating to the structural variables. This includes variables which reflect the dwelling's size, age, structure type, heating and cooling equipment, and several quality measures. Neighborhood variables are discussed in the second category of estimated coefficients. Neighborhood variables are constructed from the race of head of household, the respondent's opinion of the neighborhood, and the presence of abandoned housing, trash or litter, and convenient shopping. Central city, county and state variables are also included in this category. The third category covers the conditions of the contract for renters and flow of services for home owners. The estimated coefficients for length of tenure, persons per room, and adjustments for utilities, furniture and parking are discussed in this section. The fourth and final category includes several variables that measure inflation in housing prices.

Within each category the distributions of estimated coefficients for the renter and owner models are compared whenever possible. Finally, we suggest additional analysis of the coefficients to be performed at a later date.

The average coefficient estimates for the renter and the owner equations are presented in Exhibits 7 and 8. These exhibits display the mean and median estimated coefficient for each regressor.

A non-symmetric distribution or the presence of outliers causes the mean and median to take on different values. These exhibits also list the standard deviation and interquartile range for the distribution of each estimated coefficient.

In summary, the important features of a distribution of coefficients are the average value, the statistical significance of the estimated coefficients, the existence of outliers, and the estimated coefficients with unexpected signs. The discussion begins with the estimated coefficients for structural type variables, followed by variables which measure neighborhood effects, contract conditions, and inflation.

Dwelling Size Coefficients

The first group of structural coefficients are those related to the size of the dwelling. These variables are constructed from survey questions about the number of bathrooms, bedrooms, and other rooms; and for owners, whether the dwelling has a garage or basement. The distributions of estimated coefficients in this category are the most well behaved in both the renter and owner models. That is, they are generally symmetric about the average estimated coefficient, with almost all coefficients possessing the correct sign. The dwelling size estimates are almost always statistically significant at the 10 percent level. Recall the renter model constant term represents units having one bathroom, one bedroom, and two other rooms, so the estimated coefficients of the dwelling size variables must be interpreted as differences from this standard dwelling. The standard unit in the owner equation is any unit having one bathroom, three bedrooms, and three other rooms.

By examining the mean, or the median, estimated coefficient for the dwelling size variables, we see the advantage of using dichotomous variables to indicate the number of bathrooms, bedrooms, and other rooms. Consider the bedroom variables in the renter model. There is an average 20 percent discount going from a standard unit with one bedroom to a unit with no bedroom. However, two and three bedroom units command premiums of 12 and 23 percent over comparable one bedroom units. If the 20 percent discount for no bedroom were constant, two and three bedroom units would command premiums of 20 and 40 percent, respectively, compared to the one bedroom dwelling. A single variable for the number of bedrooms would not have captured this nonlinear effect.

The average values of the estimated coefficients for the number of bathroom variables, on the other hand, are approximately linear. In the owner equation, for example, a house with 1.5 bathrooms commands an 11 percent premium over a house with only one bathroom. A two bathroom house is worth, on average, 18 percent more than a one bathroom dwelling. The 18 percent premium is approximately one standard deviation away from a 22 percent premium--twice the premium associated with the 1.5 bathroom house. Similarly, a three (or more) bathroom house requires a 38 percent premium over a one bathroom house. Since the standard deviation for the distribution of estimated coefficients is 8.6 percent, the mean coefficient is within one standard deviation of a 44 percent premium. A 44 percent premium would be required if we used 11 percent for each additional half bath. In spite of the apparent linear relationship among the mean estimated coefficients for the number of bathrooms variables, we recommend using dichotomous variables. The approximately linear relation of the average coefficients does not apply to individual bathroom coefficients in many SMSAs. In addition, there is no need to collapse the categories to one variable to conserve the number of degrees of freedom because there are hundreds of observations in each regression.

Exhibits 7 and 8 list the number of times the estimated coefficients are statistically significant for each distribution. The number of bathrooms coefficients are always statistically significant in the owner model and nearly always significantly different from zero in the renter model. The owner model shows a similar tendency to produce more significant coefficients for the number of rooms other than bedrooms coefficient. This tendency is reversed with the number of bedrooms variables. For example, 45 of the 59 estimated coefficients corresponding to 4 bedroom owner-occupied units are statistically different from zero. All the estimated coefficients for 4 or more bedroom units are statistically significant in the renter model. The indicator variable for the presence of a garage is positive and statistically significant 57 times. The indicator variable for basement is positive and significant 42 times, and negative and statistically significant once.

An estimated coefficient is an outlier if its value is unlike the values of the other coefficients in the distribution. The discussion of outliers, unless otherwise noted, is limited to the coefficients that are statistically different from zero. Few dwelling size coefficient estimates differ radically from the norm, but there are several worth mentioning.

The estimated coefficient for the intercept for the Honolulu renter equation is much greater than the average of the intercepts for the other SMSAs. This implies the standard dwelling is much more expensive in Honolulu than anywhere else. Comparative housing cost data collected by the Bureau of Labor Statistics in 39 cities also find Honolulu to have much more expensive rental housing than other urban areas.[1] The high value of the intercept is accompanied by outlying values for several other coefficients. For example, the estimated coefficient for a unit with 4 rooms other than bedrooms in the Honolulu owner equation is the only estimated coefficient for that variable which is not significantly different from zero. Besides having an unusual concentration of outlying coefficients, the Honolulu equations have the lowest R^2 statistics for both renter and owner models.

Outlying estimates for three other dwelling size coefficients are worth mentioning. The estimated coefficient for no bedrooms in the Baltimore renter equation (-.42) is far below the next lowest estimated coefficient of Cincinnati (-.32). In the owner model, outliers appear in the distributions for the estimated coefficients of the R12 (units with one or two rooms other than bedrooms and bath-

---

[1] See the rent component in Urban Family Budgets in U.S. Department of Labor (1977).

rooms) and BEDG4 (4 or more bedrooms) variables. Washington, D.C.'s estimated coefficient for R12 is a positive outlier and the only estimated coefficient not significantly less than zero. Finally, the value of Raleigh's estimated coefficient for a 4 or more bedroom house (0.0498) is more than 3 standard deviations below the mean of the distribution.

To conclude this discussion we indicate those SMSAs with consistently high or low estimated coefficients in several dwelling size categories. The estimated coefficients for the number of bathroom variables in the renter model are consistently low in Anaheim and in Buffalo. The SMSAs with high estimated coefficients in this category are Boston, Fort Worth, and New York. The Baltimore, Albany, Columbus, Detroit, and Rochester SMSAs have consistently low estimated coefficients for the number of bedroom variables in the renter equation. Similarly, the SMSAs with consistently high estimated coefficients in this category are Anaheim and San Antonio. In the owner model, the low estimated coefficients for the number of bathroom variables correspond to the Paterson and Newark SMSAs and the high estimated coefficients are found in Wichita, Dallas, and Oklahoma City. The Honolulu SMSA has a consistently low estimated coefficient for several of the number of rooms less bedrooms variables. Finally, the Raleigh, Dallas, Oklahoma City, and Phoenix SMSAs all have high estimated coefficients for the number of rooms less bedroom variables.

The coefficient estimates for dwelling size characteristics are among the most consistent performers in the hedonic regressions. These estimates are nearly always statistically different from zero, have the expected sign, and exhibit few outliers.

Structure Type

The second category of structural coefficients are the structure type coefficients. The renter model includes the variables SFATT (single-family attached dwellings), SFDET (single-family detached dwellings), DUPLEX (two units), and NGT50 (more than 50 units). The omitted category incorporated in the intercept is a unit in a building that has between 3 and 50 units. The owner equations have only one dwelling type variable, SFATT. Single family detached dwellings is the omitted category contained in the intercept. Multiple unit dwellings are omitted from the owner hedonic because the Annual Housing Survey only reports the value of single family houses.

For renters, the coefficients of structure-type variables have signs which indicate that these variables are probably picking up locational effects rather than structural effects. The more desirable low density structure types, SFDET, SFATT and DUPLEX, have consistently negative coefficients while structures with more than 50 units have largely positive coefficients. Since larger structures are typically located in areas with greater accessibility and higher land costs, the observed pattern could well be reflecting this locational difference. For owners, the distribution of estimated coefficients for single-family attached units is approximately normally distributed around zero. The coefficient of SFATT is negative and significant in

11 SMSAs and positive and significant in 17. The opposing benefits of low density and more accessibility probably account for some of the vacillation of the coefficient between discount and premium in the owner model. Older, slow growing areas often suffer the greatest decline in their central cities while the suburbs grow. On the other hand, rapidly growing areas tend to have vibrant central cities and suburbs. In slow growing areas low density could be dominating accessibility while the reverse occurs in rapidly growing areas. The five lowest estimated coefficients for SFATT are all slower growing, northeastern SMSAs: Albany (-0.38), Baltimore (-0.33), Allentown (-0.33), Philadelphia (-0.28), and Pittsburgh (-0.17). At the other extreme, four of the five largest estimated coefficients appear in western, rapidly growing SMSAs. These SMSAs are Wichita (0.47), Spokane (0.44), Los Angeles (0.28), and Denver (0.26). Milwaukee (0.28) is the fifth SMSA in this latter group.

Age of the Structure

Housing is a durable good which takes many years to be completely consumed. Some of the unit is consumed or used each year as paint and wood age and roofs wear down. If all units are identically constructed, inflation is absent, and the rate of maintenance and repair expenditures is the same for all units, then precise measurement of the rate of depreciation is possible by observing the value or rent of two or more units of different ages. This is not possible, however, because inflation does exist; because units are constructed differently; and because some households spend more on maintenance, repair and alterations than others. In order to estimate accurately the effect of aging on values and rents, it is necessary to control for inflation and quality differences in housing units. The hedonic technique is one way to control for differences in dwelling quality and inflation rates, but it cannot control for most differences in maintenance. The hedonic equations yield estimates of how rents and values depend on the inexorable ravages of time, and maintenance decisions made to combat that decay.

What is the importance of obtaining an accurate estimate of the depreciation rate? First, depreciation is a cost so that the faster dwellings depreciate the higher is the cost of housing. Second, because it is a cost, depreciation is an allowable deduction in computing taxable income from rental housing. Current tax law allows depreciation rates of 3 to 6 percent in early years of a project's life.[1] If true depreciation rates are less than those allowed for tax purposes, this is an important incentive for expanding the supply of rental housing.[2] However, there is little evidence on true depreciation rates so the size of the stimulus, if any, is hard to gauge. The hedonic estimates of depreciation presented here can help in quantifying that stimulus.

Age Variable Specification and Interpretation

The measure of depreciation in the hedonic regressions is derived from several variables measuring the structure's age. AGE1 is the age of the structure, and is constructed from a survey question about

---

[1] Straight line depreciation on 31 years of useful life gives 3.2 percent annual depreciation. Double declining balance depreciation on a 31 year project gives 6.4 percent depreciation the first year, 6.0 percent in the second year, and so on.

[2] See deLeeuw and Ozanne (1980), and Wykoff (undated).

when the house was built, and the date of the AHS Interview. The coefficient of AGE1 can be interpreted as the percentage change in rent or value given a one year change in age. However, it is not known that the depreciation rate for a dwelling will be the same when it is new as when it is 30 years old. Certainly automobiles depreciate faster in the first year than in later ones. To avoid constraining the depreciation estimate to be constant we include higher order terms AGE1SQ (age squared) and AGE1CB (age cubed).

Another problem involves measuring the age of very old dwellings. The AHS survey question on the date the structure was built asks the year built for post-1970 dwellings; earlier years are collapsed into six categories (1969-1970, 1965-1968, 1960-1964, 1950-1959, 1940-1949, and pre-1940). The mid-year of these intervals is used to construct the age of five of the six cohorts, but it is difficult to assign a reasonable number to the pre-1940 cohort. This category is open ended, and the average age of structures in this category probably varies greatly from city to city. For example, San Francisco has a fair number of turn-of-the-century dwellings in its current housing stock; Fort Worth has few, if any. The variable DAGE is an indicator variable signifying that a dwelling is in the early cohort. This variable allows estimation of depreciation rates in earlier years to remain unbiased by city-to-city differences in this oldest cohort.

AGE1, AGE1SQ, and DAGE are included in the renter equations; AGE1CB is added to these in the owner regressions (the cubic term was always statistically insignificant in preliminary renter regressions). The discussion which follows revolves around the continuous variables AGE1, AGE1SQ and AGE1CB. The coefficient estimates for DAGE are presented but not interpreted.

The estimated coefficients of the age variables are difficult to interpret individually. To properly explain the effect of age on rent or house value, the effects of all continuous age variables must be considered simultaneously. For example, if the estimated coefficients for AGE1, AGE1SQ, and AGE1CB are -.01, +.001 and -.0001, respectively, the estimated discount of a three year old dwelling relative to a new one is 3(-.01) + 9(.001) + 27(-.0001), or a discount of 2.37 percent. The price of a four year old dwelling is 4(-.01) + 16(.001) + 64(-.0001), or a 3.04 percent discount. The estimated depreciation rate <u>at three years</u> is the percentage decrease in value or rent at that time. Consider the age terms of the semi-logarithmic hedonic model:

$$\text{(A-1)} \quad \ln V = \hat{\beta}_1 A + \hat{\beta}_2 A^2 + \hat{\beta}_3 A^3 + \hat{\alpha} X + e$$

where V is rent or value,

$\hat{\beta}_i$ and $\hat{\alpha}$ are estimated coefficients,

A is the age of the structure,

X represents all other independent variables, and

e is the residual.

Taking the derivative with respect to age gives the depreciation rate

$$\text{(A-2)} \quad \frac{dV/dA}{V} = \hat{\beta}_1 + 2\hat{\beta}_2 A + 3\hat{\beta}_3 A^2$$

Using the numbers from our example above, the estimated depreciation rate at three years is:

$(-.01) + 2(.001)(3) + 3(-.0001)(9) = -.67$ percent

Note that the percentage decrease from year three to year four, 2.37 percent minus 3.04 percent, is also -.67 percent. This simple difference is a very good approximation to the true depreciation rate for the values of age we deal with. Depreciation rates reported in this section will be calculated from equation (A-2), however.

Estimation Results

Each of the distributions for the age of structure variables have statistically significant negative and positive coefficients. The signs for the coefficients of AGE1, AGE1SQ, AGE1CB, and DAGE typically oscillate for each SMSA. That is, when the coefficient of AGE1 is positive, the coefficients of AGE1SQ and DAGE are negative. Similarly, when the coefficient of AGE1 is negative, the coefficient of AGE1CB is negative and the coefficients of AGE1SQ and DAGE are positive. What do all these numbers mean? One way to effectively summarize the results is to plot the estimated discounts in rent and value by the age of the dwelling. This was done for all 118 models. All the results are not presented here for lack of space, but Exhibit 9 presents several such plots. The first two plots represent typical results. They are the mean discounts in rent and value by age for the 59 SMSAs. The other plots are less representative of the typical city but include features of interest we will discuss below.

Another summary measure of these results is presented in Table 3. These are depreciation rates for renters and owners at selected ages. Columns 1 and 4 give the mean for all SMSAs, the other columns are for the SMSAs named. The averages represent more SMSAs than the selected SMSAs in Table 3, which were chosen to illustrate patterns different from the average.

Do not be confused by the difference between the plots and the table. The plots are of discounts, or the _level_ of the implicit price of age, while the table presents depreciation rates, which are the _rates of change_ in the price of age. The former tells what a dwelling is worth, relative to a new dwelling, while the latter measures how fast its worth is changing with increasing age. The following summarizes our depreciation estimates.

Rents and values consistently decrease with age. This is, of course, intuitively appealing. On the average, a two year old dwelling rents for one percent less than a new dwelling, while a two year

Table 3

Depreciation Rates for Owners and Renters,
for Selected Years, Selected SMSAs*

	Renters			Owners		
	Average	Boston	Anaheim	Average	Miami	San Diego
Rate at Year 1	-.0060	.0112	-.0173	-.0092	.0168	-.0033
Rate at Year 2	-.0060	.0104	-.0165	-.0084	.0147	-.0027
Rate at Year 9	-.0059	.0055	-.0108	-.0045	.0031	.0010
Rate at Year 10	-.0059	.0048	-.0100	-.0042	.0019	.0015
Rate at Year 20	-.0059	-.0024	-.0019	-.0029	-.0040	.0052

*SMSAs are selected to show rates significantly different from averages.

old owner-occupied dwelling is worth two percent less than when new. The average discounts for renters and owners can be read from the plots in Exhibit 9.

There are several estimates of appreciation and of essentially zero depreciation. Miami owners and Boston renters pay up to a 10 percent premium for a ten year old dwelling relative to a new one. Detroit renters pay premiums for older dwellings in the first twenty years; past that older dwellings are heavily discounted. Allentown renters pay up to an 11 percent premium for a 16 year old dwelling. The estimates of San Diego and Philadelphia owner depreciation rates, as well as that for Cleveland renters, are statistically zero. That is, age has no measured net effect on rents or values in those SMSAs.

Depreciation rates, and discounts for age, differ among SMSAs. The averages adequately represent a fair proportion of the estimates, but there are several cities which exhibit differences from the average pattern. Future work will explain systematic differences among estimated depreciation rates.

The relative prices of older dwellings differ between tenure groups, even in the same SMSA. This is not surprising. For example, consider that rents are returns to the current flow of housing services, while values are a stock concept. Rents might be expected to change mainly as the flow of housing services decreases as the units deteriorate. Values, on the other hand, reflect the present value of future housing services, as well as current services. Thus present value changes with the expected life of the house as well as with the change in the current flow of services as the unit ages.

Exhibit 9 demonstrates that, on average, owner dwellings depreciate faster in the first five years of a dwelling's life, but that for the next 25 years rents drop faster.

On average, rents decrease at a nearly constant rate, values at a declining rate. The average depreciation rate for renters is remarkably constant, ranging from 0.58 to 0.60 percent. Owner depreciation rates show more variation, from 0.9 percent in year 1 to 0.28 percent in year 20. By year 30, the only observation point after year 20 for measuring depreciation, the rate rises again to 0.6 percent. These depreciation rates are similar to the 0.5 percent estimate of Frank C. Wykoff (undated) for the value of rental properties. He used a hedonic approach similar to ours but lacking as much detail. Depreciation estimates by other methods and from other assets, e.g., commercial buildings, generally find higher depreciation rates, such as 1.0 to 2.5 percent.[1] The difference may be due in part to better controls on physical features of the dwelling. In any event, all hedonic depreciation rates suggest that current tax depreciation schedules considerably overstate true depreciation.

A Caveat. So far in this discussion, the coefficient of our age variables have been interpreted as an accurate indicator of depreciation. It is likely, however, that the model is imperfectly specified. For one example, the AHS lacks some desirable locational information such as distance to the central business district (CBD). Thus we may not be capturing all the influences of location upon dwelling rent or value. If so, and if age is correlated with

---

[1] Bruggeman (1977), Hulton and Wykoff (1978), Taubman and Rasche (1969), Palmquist (forthcoming).

location--as it seems to be--then the coefficients of AGE may be biased. Of course, the direction of the bias depends on whether the older dwellings are located in the more or the less desirable locations.

Another source of bias is that old units which have dropped out of the stock are, of course, excluded from the sample. That is, sixty year old units are included only if they still command a positive rent or value. Sixty year old units which no longer command a value are excluded. This procedure, although it is unavoidable, does produce estimates which understate the average rate of depreciation for all 60 year old dwellings.

Changing construction quality can introduce another bias into the hedonic estimates of depreciation. If in the 1940s and 1950s dwellings were built with higher construction quality than those built in the 1960s and 1970s, then these old units will not have fallen in value relative to new units by as much as they have relative to their replacement cost. Since hedonic equations cannot control for construction quality very well they will understate depreciation in such cases. The reverse bias occurs if older dwellings are of lower construction quality relative to new dwellings.

In which direction is the net bias of our results? It is impossible to know. Our estimates tend toward the lower side of the previous estimates. However, none of the small number of depreciation studies is sufficiently definitive to draw firm conclusions. This must await further study.

Dwelling Equipment

The fourth category of structural coefficients are the estimates for dwelling equipment. These coefficients describe the heating system, cooling system, and for renters of dwellings in multiple unit buildings, whether the building has an elevator. The variables that describe the heating system for renter are RHEAT, an indicator variable for a unit having wall or room heaters with flue, and POOR, a linear combination which includes an indicator variable for primitive heating equipment such as portable heaters. The renter equation includes POOR and RHEAT. The owner equation includes RHEAT, SHEAT, and EHEAT. SHEAT indicates a house heated with steam or hot water heat, and EHEAT indicates a unit heated with electricity. Note that the owner indicator variables are not mutually exclusive because there are both equipment variables and fuel variables. For example, one can have an electric room heater, in which case EHEAT and RHEAT both take the value one. The omitted heating equipment category is central warm air heating for owners and central warm air, steam or hot water for renters. The cooling equipment variables are ROOMAC, an indicator variable for room air conditioners, and CENTAC, an indicator variable for central air conditioning. Finally, ELEVP, an indicator variable for units that are serviced by an elevator, is included in the renter equation.

The average estimated coefficients for the dwelling equipment variables all have the correct sign. For example, the estimated coefficient for RHEAT is negative and statistically significant in 56 SMSAs in the renter model and in 38 SMSAs in the owner model. The estimated coefficient for RHEAT is never positive and statistically significant.

The estimated coefficients for air conditioning enjoy similar success. The estimated coefficient of ROOMAC is positive and statistically significant 48 times in the renter equation and 29 times in the owner equation. This estimated coefficient is negative and statistically significant once in the renter equation (San Francisco) and four times in the owner equation. The estimated coefficient of CENTAC is positive and significant 53 times and 54 times in the renter and owner equations, respectively. The largest average premium or discount associated with dwelling equipment variables in the renter and owner models is the coefficient for CENTAC. The presence of central air-conditioning requires an average premium of 19 percent in the renter equation and 13 percent in the owner equation.

The outlying estimated coefficients for dwelling equipment variables are the coefficients of RHEAT in Honolulu renter (-0.54) and owner (0.47) equations and the coefficient of CENTAC in the Tacoma renter equation (-0.27). The largest estimated coefficients in the owner equation for the airconditioning variables both occurred in Houston (a 22 percent premium for room airconditioning and a 31 percent premium for central airconditioning). In the renter equation, New Orleans has the largest estimated coefficient for ROOMAC (0.19) and the second largest estimated coefficient for CENTAC (0.34). Other SMSAs with high estimated coefficients for the airconditioning variables are Memphis and San Antonio in the renter equation and Dallas and New Orleans in the owner equation. All these large coefficients occur in the South where airconditioning is understandably more important.