Working Paper 85-6

LOGIT ANALYSIS OF THE EFFECT OF RENT CONTROL ON HOUSING QUALITY

by

David L. Mengle*

Federal Reserve Bank of Richmond

July 1985

*The views expressed in this paper are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

I. Introduction

Rent control is one of the few policy issues on which there is general agreement among economists.[1] Economic theory predicts, and few economists have tried to dispute, that imposing rent controls on a housing market is likely to lead to rental housing shortages and general deterioration of quality. Even on income distribution grounds, rent control receives poor reviews, since it is generally agreed to be a clumsy and inexact means of helping the poor.

Despite this seeming consensus, however, economists have apparently had little success in influencing policy discussions. During the accelerating inflation of the 1970's, rent controls were enacted in Massachusetts, New Jersey, and California, along with Washington, D.C. The slackening of interest in controls since that time is probably due more to the decline of inflation and rental housing demand than to any evidence advanced by researchers on the laws' ill effects.

An interesting aspect of the debate during the 1970's was that rent control proponents claimed that the laws they favored, known as "second generation" rent controls, were designed to avoid the deterioration and abandonment often attributed to the older controls in New York City. The newer laws contained mechanisms for controlled increases and passing some operational cost increases through to tenants, along with penalties for allowing housing quality to fall. Such provisions, it was argued, would protect tenants from rent gouging while encouraging landlords to maintain their properties through allowing a "fair" return and denying rent increases to deteriorated dwellings.[2] In fact, one rent control advocate cited evidence he claimed demonstrated that rent control actually led to improved housing quality.[3]

---

[1] J. R. Kearl _et_ _al_. (1979).

[2] See Richard E. Blumberg _et_ _al_. (1974).

[3] John I. Gilderbloom (1980).

Most economic analyses of the effects of rent control on housing quality have been theoretical in nature.[4] Essentially, these studies predict that, since maintenance expenditures are one of the few factors of production of housing services that are variable in the short run, landlords will respond to lower rents (in real terms) by reducing maintenance and, therefore, quality. In addition, since the effects of lower maintenance expenditures are cumulative in nature, it is reasonable to expect that such deterioration will increase over time as controls become more binding. Despite these plausible theoretical predictions, however, there is very little empirical evidence to support such contentions. Reasons for this gap include little experience, outside New York, with rent controls, along with a lack of detailed quality data that has only begun to be remedied in the last decade.

The purpose of this paper is to look for evidence that rent controls lead to lower housing quality than would have been the case had controls not been enacted and that such effects increase over time. In the following section, the hypotheses to be tested and the data to be used will be described, followed by a derivation of the empirical models to be used and a presentation of the results of the estimated models. Instead of attempting to develop and use a measure of quality, the analysis will simply look for a systematic relationship between the existence of rent control and other factors and whether or not a dwelling is judged to be of satisfactory quality according to a set of governmental criteria. The final section will comment on the policy implications of the results.

---

[4] See, for example, John C. Moorhouse (1972) and Mark Frankena (1975).

II. Empirical Analysis

A. Statement of hypotheses

In this paper, two hypotheses about rent control and housing quality will be tested. The first is common sense and is implied by theoretical models of rent control found in the literature. This hypothesis is that landlords in controlled markets will cut back maintenance expenditures in order to maximize profits (or minimize losses), thereby causing quality to deteriorate to a lower level than would have been the case in a free market.

The second hypothesis to be tested is that the effects of rent controls will become more pronounced over time. The rationale for this is that disinvestment in response to controls will not have immediate pronounced effects, but rather will manifest itself slowly through depreciation.[5]

In order to test these two hypotheses, it is crucial that a usable method of evaluating quality be developed. In the following section, the data will be described, along with the quality measure to be used.

B. Data

The hypotheses described above will be tested using the Annual Housing Survey conducted by the United States Bureau of the Census in cooperation with the Department of Housing and Urban Development. This data set helps to fill in part the severe gap that previously existed between the requirements of economic theories of housing markets and the actual information to be had. The Annual Housing Survey has been conducted since 1973, although usable data for time series actually begin in 1974. Over a three-year period, sixty Standard Metropolitan Statistical Areas (SMSA) are sampled, with the same

---

[5]See C. P. Rydell _et al_. (1981), pp. 55-9.

one-third being visited every three years. Further, the same households are visited in each "wave," thereby giving a longitudinal character to the survey.

This analysis will use a combined data set of 8,281 dwellings from the central city portions of eight SMSAs, of which four are controlled and four are not. Size and control status are summarized in Table 1. This set of 8,281 dwelling units is selected according to several criteria. All rental units that were added to or withdrawn from the market during the period of observation are excluded from the sample. All dwelling units selected must be private, unsubsidized rental units. In addition, single family dwellings along with two-family detached units are excluded since they would be exempt from controls in most of the cities in the sample.

Observations on each of the dwellings in all eight SMSAs will be pooled into one sample under the assumption that a "national" (or, at least, a northeast-north central regional) housing market exists. This may be justified in two ways. First, variables will be included to pick up differences between cities. Second, Peter Linneman's (1980) tests of the national housing market hypothesis show that the functional form for hedonic rent equations is generally the same for his analyses of individual cities (Chicago and Los Angeles) as for his pooled sample of 34 major American cities. Further, the differences between coefficient estimates for the individual cities and the pooled sample are insufficient to reject the national housing market hypothesis.

The AHS contains detailed information on the characteristics of dwellings, including data on neighborhood quality, dwelling quality, structure, costs, geography, previous residence, and characteristics of the occupants. Variables to be used in this analysis are listed in Table 2. Unfortunately, this set contains no data on landlord maintenance expenditures. Therefore, it is necessary to look for signs that point to reduced maintenance. This would

TABLE 1

CITIES IN POOLED DATA SET

City	N	Rent Control Law?
Boston	2,086	Yes
Detroit	844	No
Minneapolis-St. Paul	484	No
Newark	291	Yes
Paterson-Clifton-Passaic	377	Yes
Philadelphia	1,375	No
Pittsburgh	185	No
Washington	2,639	Yes
Total	8,281	

Observations for Paterson-Clifton-Passaic and Philadelphia were made during 1975 and 1978; all others are from 1974 and 1977.

TABLE 2

VARIABLE DESCRIPTIONS

	Dwelling Characteristics (z)
BUILT	Age of building
OWNHERE	Owner dummy = 1 if owner lives in building
DUMP	Neighborhood quality variable = 1 if there are no rundown houses or buildings in neighborhood
RACE	Neighborhood race variable = 1 if head of household is nonwhite
	Market Characteristics (x)
RC	Law variable = 1 if city had a rent control law during 1974-77 period. Source: Lett (1976), pp. 72-6.
VAC	City vacancy rate. Source: U.S. Bureau of Census, Annual Housing Survey: Housing Characteristics for Selected Metropolitan Areas. Various issues, 1977, 1978.
YPC	Per capita income. Source: International City Management Association, Municipal Year Book: 1981.
	Interaction Terms (w)
RBUILT	(RC) x (BUILT)
ROWNH	(RC) x (OWNHERE)
RDUMP	(RC) x (DUMP)
RRACE	(RC) x (RACE)

Unless otherwise noted, variables are from Annual Housing Survey tapes, 1974-7 and 1975-8 linked samples.

mean attempting to isolate those variables which depend on the landlord's actions. These may include water breakdowns, toilet breakdowns, holes in floors, cracks in the walls and ceiling, rats and mice, and others. In addition, data are available on the number and duration of some of these faults in a building.

The most important task for any investigation of the effect of rent control on housing quality is to develop a usable measure of quality. To begin with, it is reasonable, as pointed out in the previous paragraph, to focus only on those characteristics of a dwelling that may be altered by landlords at relatively low cost. It is not obvious, however, how these variables should be aggregated into a scalar measure. One possibility is weighted aggregation of individual variables measuring attributes that can be affected by a landlord's choice of maintenance expenditures.[6] Here, the indication of relative importance of the various faults is either arbitrary or could, conceivably, be based on coefficients from regressions of rent on characteristics. Unfortunately, such coefficients are often of low magnitude and no statistical significance, if not with signs opposite to what one may reasonably expect. Another possibility is factor analysis or principal components regression, both of which allow the data to mechanically show the way to aggregation. Unfortunately, these tools give little help in the way of economic interpretation. Thus, it becomes necessary to look for a third alternative.

In this study, quality of dwellings will be rated either "satisfactory" or "unsatisfactory" based on a Congressional Budget Office report on federal

---

[6] For illustrations of the difficulties of formulating indicators of the level of housing quality, see Jeanne E. Goedert and John L. Goodman, Jr. (1977) and Mengle (1983), pp. 88-94.

housing policy. This report makes use of AHS data, and develops a set of criteria for classifying a dwelling as "in need of rehabilitation." The criteria are the following:

> A unit was classified as needing rehabilitation if it had at least one of the following conditions: (1) the absence of complete plumbing facilities; (2) the absence of complete kitchen facilities; (3) the absence of either a public sewer connection, a septic tank, or cesspool; (4) three or more breakdowns of six or more hours each time in the sewer, septic tank, or cesspool during the prior 90 days; (5) three or more breakdowns of six or more hours each time in the heating system during the last winter; (6) three or more times completely without water for six or more hours each time during the prior 90 days; (7) three or more times completely without a flush toilet for six or more hours each time during the prior 90 days; and/or if the unit had two or more of the following conditions: (1) leaking roof; (2) holes in ceiling; (3) open cracks or holes in interior walls or ceilings; (4) broken plaster over greater than one square foot of interior walls or ceilings; (5) unconcealed wiring; (6) the absence of any working light in public hallways for multi-unit structures; (7) loose or no handrails in public hallways in multi-unit structures; (8) loose, broken, or missing steps in public hallways in multi-unit structures.[7]

In terms of the analysis to be conducted here, several of the above conditions will not be of particular interest because they are not easily altered by the landlord. This refers to such characteristics as plumbing facilities and sewer connections. In addition, unconcealed wiring may be easily removed but it is difficult to understand why wiring would be exposed in response to lower rents. Finally, blown fuses, peeling paint, and evidence of rats and mice will be added to the above criteria because they are easily affected by a landlord's actions. The revised set of criteria is shown in Table 3, and the resulting quality variable (d_i) will take a value of one if a dwelling is rated satisfactory and zero if unsatisfactory.

---

[7]U.S. Congress, Congressional Budget Office (1978), p. 6.

TABLE 3

CRITERIA FOR UNSATISFACTORY CLASSIFICATION

ANNUAL HOUSING SURVEY

1. One or more of the following occurring three or more times each:

   a. No running water for 6 hours or more in last 90 days
   b. Toilet breakdowns of 6 hours or more in last 90 days
   c. Heating breakdowns of 6 hours or more last winter
   d. Blown fuses or breakers in last 90 days

2. Two or more of the following:

   a. Roof leaks
   b. Open cracks or holes in walls or ceilings
   c. Holes in floor
   d. Peeling paint or broken plaster over one square foot (1974-5 only)
   e. Peeling paint over one square foot (1977-8 only)
   f. Broken plaster over one square foot (1977-8 only)
   g. All or some public light fixtures not working
   h. Hazardous steps on common stairway
   i. Stair railings missing or not firmly attached
   j. Signs of rats or mice in last 90 days

C. Empirical models

In this section, two different empirical models will be presented, a linear probability model and a logit model. In both, it will be assumed that there is an unobservable vector (z_i) of alterable quality characteristics that is positively related to a landlord's maintenance expenditures for a dwelling i. To the analyst, a certain level of z_i will manifest itself in more or less frequent maintenance problems of various sorts. In other words, if certain defects are observed during a given period, it will be taken as an indication of a relatively high probability of a low level of the unobservable z_i. For example, cracks in walls or ceilings, more than a certain number of system breakdowns, leaks in the roof, or loose railings point to the possibility of lowered maintenance. Although these defects can certainly occur in normally well maintained units, they are more likely to occur if the unit is poorly maintained.[8] Thus, dwellings will be simply classified as satisfactory or unsatisfactory as a reflection of the maintenance they receive from their owners or managers.

In order to make this operational, a qualitative response model will be developed. Assume that z_i can be classified into two categories, satisfactory and unsatisfactory. Assume further that z^* is the threshhold level of z_i below which a dwelling is classified unsatisfactory. Thus, let

$$d_i = \begin{cases} 1 & \text{if } z_i > z^* \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The distribution of quality over satisfactory and unsatisfactory is shown in Figure 1. What is of interest here is the probability that a dwelling will be classified satisfactory, which will be called P(d=1).

---

[8] Kenneth F. Wieand (1983) also emphasizes the relationship between the quality of a dwelling and the probability of observing a particular defect.

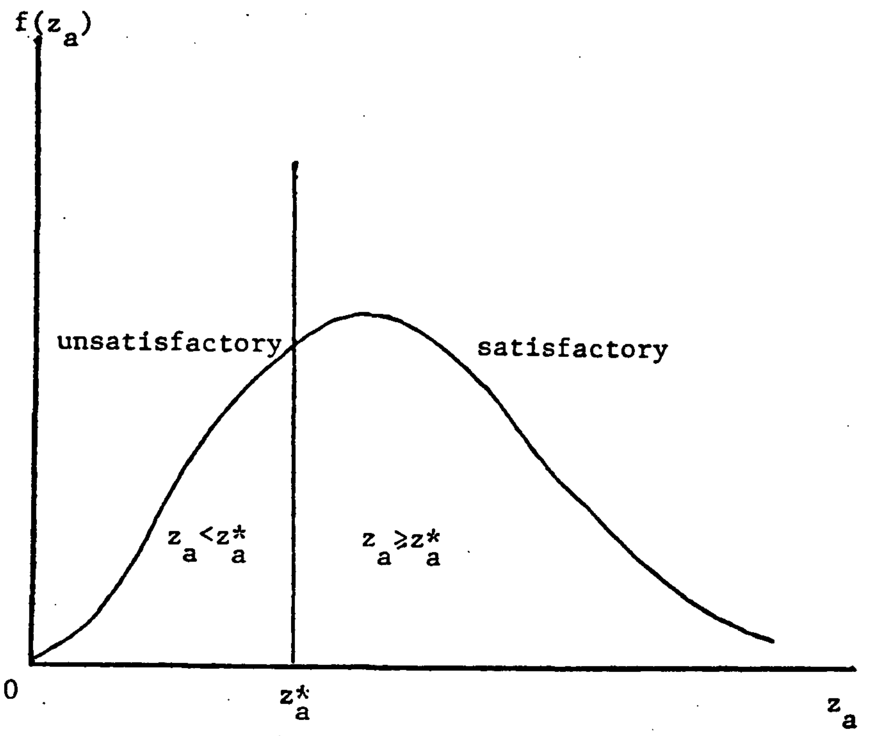

Figure 1: Distribution of Quality

The dichotomous variable (d_i) will take the value one if a dwelling is rated satisfactory, and zero otherwise. An unsatisfactory rating is assigned if at least one deficiency under the first heading of Table 3 occurs, or at least two under the second heading occur, or both.[9]

---

[9]Orhan Yildez (1983) uses a criterion that rates a dwelling unsatisfactory if any deficiency is found.

The dichotomous variable will be used in two ways. First, this study will use a linear probability model, in which the left hand side variable is the probability that a dwelling is rated satisfactory. Thus, the model to be estimated is the following:

$$d_i = \alpha^p + \beta_1^p z_{1i} + \ldots + \beta_4^p z_{4i} + \gamma_1^p x_{1i} + \ldots + \gamma_3^p x_{3i} + \tau_1^p w_{1i} + \ldots + \tau_4^p w_{4i} \qquad (2)$$

where the "p" superscripts designate the linear probability model parameters, and the z_i variables refer to dwelling i characteristics, the x_i refer to the city in which dwelling i is located, and the w_i are interaction terms between the existence of rent control and dwelling traits. Specifically:

d_i = dichotomous quality variable

z_1 = age of building (BUILT)

z_2 = presence of owner in building (OWNHERE)

z_3 = neighborhood variable (DUMP)

z_4 = race of head of household (RACE), used for neighborhood race

x_1 = presence of rent control law (yes = 1, no = 0)(RC)

x_2 = city vacancy rate (VAC)

Source: Current Housing Reports: Annual Housing Survey. Various issues, 1977, 1978.

x_3 = per capita income, 1977 (YPC)

Source: ICMA, Municipal Year Book: 1981.

w_1 = Interaction: Rent control with age (RBUILT)

w_2 = Interaction: Rent control with presence of owner (ROWNH)

w_3 = Interaction: Rent control with neighborhood quality (RDUMP)

w_4 = Interaction: Rent control with neighborhood race (RRACE)

Unfortunately, there are severe problems connected with the use of the linear probability specification. To begin with, consider the model:

$$d_i = \alpha + \beta z_i + \varepsilon_i \qquad (3)$$

the fact that d_i can only take on values of zero or one and that $\alpha + \beta z_i$ can take on any value means that ε_i can take on only one of two values, $-(\alpha + \beta x_i)$ if $d_i = 0$ and $1 - (\alpha + \beta x_i)$ if $d_i = 1$. Since the distribution of ε_i must be the same as that of d_i when z_i is assumed to be nonstochastic, it follows that ε_i follows a binomial distribution. Assuming $E(\varepsilon_i) = 0$, the probabilities required to satisfy this assumption are $1 - (\alpha + \beta z_i)$ for $d_i = 0$ for $d_i = 0$ and $(\alpha + \beta z_i)$ for $d_i = 1$. Thus, the variance of ε_i is

$$\mathrm{Var}(\varepsilon_i) = (\alpha + \beta z_i)[1 - (\alpha + \beta z_i)] \qquad (4)$$

which can be seen to be a function of z_i. This implies heteroskedasticity of ε_i and therefore unbiased and consistent but inefficient parameter estimates. In addition, estimates of standard errors of the parameter will be biased. This can be remedied by using a weighted least squares procedure, which will lead to efficient estimators. Here, ordinary least squares is used in the first stage to obtain an estimated variance-covariance matrix, which is used in the weighted least squares regression. However, Thomas A. Domencich and Daniel McFadden (1975) do not recommend weighted least squares for two reasons. First, although large sample estimates are efficient, too much weight is placed on extreme observations in small samples. Second, and of potentially greater concern in this analysis, this two stage procedure shows great sensitivity to specification error.[10]

A further problem with linear probability models is that predicted probabilities may end up outside the [0,1] interval. A possible means of

---

[10] Domencich and McFadden (1975), pp. 102-3.

attacking this problem is inequality-constrained least squares.[11] Here, β is estimated subject to the constraint that $0 < \beta z_i < 1$, which leads to a nonlinear programming problem. Domencich and McFadden point out that, although this leads to consistent estimators with lower variances, these estimates are biased. In addition, this procedure is very sensitive to specification error, which is of great concern when using a dataset like the AHS and others not likely to contain all relevant variables for the hedonic method. Given the cost of using the inequality-constrained least squares procedure, there is little to recommend it.

Thus, although there are problems in using a linear probability model, there are no good substitutes with the same simplicity of computation. It appears that the best course is to follow Takeshi Amemiya (1981) who concludes that, in the preliminary exploratory phases of a study, linear probability models are acceptable when the relative costs and benefits are considered. However, he does not recommend them for the final stage. In this study, results from all models will be reported.

The second possibility for using the dichotomous quality variable involves a logit specification, and is derived from the landlord's profit function. Although the derivation of this specification is different from that for the linear probability model, the results of the two should be consistent with each other.

The salient characteristic of this model is the emphasis it places on the results of the landlord's choices with regard to maintenance. The decision to allow a dwelling to deteriorate at a faster rate than before will be the result of its being the most profitable course of action.

---

[11] See George G. Judge, William E. Griffiths, R. Carter Hill, and Tsoung-Chou Lee (1980), pp. 588-91.

Alternatively, it will be the choice which places the landlord on the highest profit function, which is in turn assumed to result from the landlord's profit maximizing input decision. Denoting the profit function as π_{id}, and assuming each landlord owns one unit, the landlord's quality decision may be stated formally as follows. First, assume the landlord's profit function is:

$$\pi_{id} = \alpha_d^L + \beta_d^{L}{'}z_i + \gamma_d^{L}{'}x_i + \tau_d^{L}{'}w_i \qquad (5)$$

where the "L" superscript will denote the relation to the logit form, and where subscript d refers to the state of maintenance of dwelling i, and the z_i, x_i, and w_i have the same interpretation as in the linear probability model.

In order to determine the different profitabilities of maintaining and undermaintaining under the two different legal environments, the above equation may be specified as follows:

$$\pi_{i0} = \alpha_0^L + \beta_0^{L}{'}z_i + \gamma_0^{L}{'}x_i + \tau_0^{L}{'}w_i \qquad (6)$$

$$\pi_{i1} = \alpha_1^L + \beta_1^{L}{'}z_i + \gamma_1^{L}{'}x_i + \tau_1^{L}{'}w_i \qquad (7)$$

where the subscripts i refer to individual dwellings and 0 and 1 refer to choices of low or high maintenance levels, respectively. Here, π_{i1} is the profit from dwelling i if maintenance is kept at a relatively high level, while π_{i0} is the profit if maintenance is lowered. In other words, the d subscripts indicate the state of the world of high or low maintenance. The regression coefficients describe the effects of the independent variables in the two states. For example, β_0^L may be interpreted as the change in profits from undermaintenance due to a change in the city environment, while β_1^L would show the change in profits from a maintained dwelling due to a city change. With regard to the legal environment, if x_i equals zero or one depending on whether rents are uncontrolled or controlled, then the β_1^L or β_0^L coefficient would describe the effect of a change of legal environment on profits in the two possible maintenance states. A landlord will reduce maintenance if:

$$\pi_{i1} - \pi_{i0} = (\alpha_1^L - \alpha_0^L)' + (\beta_1^L - \beta_0^L)'z_i + (\gamma_1^L - \gamma_0^L)'x_i + (\tau_1^L - \tau_0^L)'w_i \tag{8}$$

$$= \alpha^{L\prime} + \beta^{L\prime}z_i + \gamma^{L\prime}x_i + \tau^{L\prime}w_i < 0$$

where the coefficients without subscripts represent the corresponding quantities in parentheses. The condition in equation (8) simply states that undermaintenance will take place if it is more profitable than maintenance.

The coefficients in equation (8) are of particular interest because they show the change in differential profitabilities between the two states of upkeep with respect to changes in the independent variables. For example, $\beta = (\beta_1^L - \beta_0^L)$ may be interpreted as the difference between effects of changes in city characteristics on the profitability of maintenance in the two states. Thus, if x = 1 if rents are controlled and x = 0 otherwise, a positive value for β_0^L means that rent control tends to make undermaintenance profitable compared to the case in an uncontrolled market. If, on the other hand, rent controls exert a negative effect on the profitability of a maintained dwelling, then β_1^L will have a negative sign. As a result, the β^L coefficient will take a negative sign in (8), signifying positive profitability for a reduced-maintenance strategy under rent control relative to what would have been the case for a free market.

Defining

$$\tilde{\pi}_{id} = \pi_{id} + \varepsilon_{id} \tag{9}$$

as the stochastic profit function, where ε_{id} is a random error term that captures influences outside the profit function specified, the probability of undermaintenance of dwelling i, that is, that $d_i = 1$, is:

$$P(d_i = 1) = P(\tilde{\pi}_{i0} > \tilde{\pi}_{i1}) \quad (10)$$

$$= P(\pi_{i0} + \varepsilon_{i0} > \pi_{i1} + \varepsilon_{i1}) = P(\varepsilon_{i1} - \varepsilon_{i0} < \pi_{i0} - \pi_{i1})$$

$$= F(\pi_{i0} - \pi_{i1})$$

where F( ) is the cumulative distribution function for $(\varepsilon_{i1} - \varepsilon_{i0})$. Assuming the $(\varepsilon_{i1} - \varepsilon_{i0})$ terms are distributed according to a logistic distribution, and omitting the superscripts to avoid clutter, the model to be estimated becomes:

$$P(d_i = 1) = \{1 + \exp[-(\alpha + \beta' x_i + \gamma' y_i + \tau' z_i)]\}^{-1} \quad (11)$$

In practice, the estimated logit model resembles the linear probability model on the right hand side. The difference is that the dependent variable is the log odds ratio $p_i/(1 - p_i)$, where p_i is the probability that dwelling i will be rated satisfactory. In the estimation of this model, maximum likelihood estimates of the parameters are obtained by using the Newton-Raphson method.

D. Results

In both models, one would expect similar signs of the coefficients. Age of building (BUILT) may be expected to have a negative sign since an older building should show more problems than a relatively new building. Second, presence of owner (OWNHERE) should show a positive sign, since in such a case the owner would be more directly affected by changes in quality. Third, the neighborhood variable, which reflects the presence of dilapidated housing near the observed dwelling, should show a positive influence on quality.

Finally, race of head of household (RACE) will serve as a neighborhood variable under the assumption of residential segregation. Since a higher value means a nonwhite head of household, the presence of discrimination means one would expect a negative sign.

Of city characteristics, the most important is the presence of rent control (RC) during the period of observation. If the hypotheses proposed above are correct, RC should show a negative sign and the magnitude of this coefficient in absolute terms should increase over time. Note that it will be assumed that laws in the cities covered by this analysis create similar incentives and will therefore be assumed to be similar across cities. The other two city variables, vacancy rate (VAC) and per capita income (YPC), are included to pick up other differences across cities.

Interaction terms are included because it is plausible that rent control laws may influence the effect of the other variables such as age and neighborhood. Since one would expect controls to aggravate the effects of other variables on quality, all the interaction terms should carry negative signs.

The results of the estimated models for the full sample are shown in Table 4. In general, the results are consistent with expectations. However, the linear probability model yields a negative but statistically insignificant coefficient for the rent control variable in 1974. However, this coefficient is significant at the 5% level using the logit specification. By 1977, as before, the rent control coefficient is significant and of greater magnitude than in 1974. Thus, according to the linear probability model, the presence of rent control reduces the probability of a satisfactory rating by nine percentage points in 1974, and by thirty-four percentage points in 1977. As shown in Table 5, rent control and its interactions are jointly significant, as are the interactions taken by themselves. Additionally, the increase in

TABLE 4
REGRESSION RESULTS: DICHOTOMOUS QUALITY VARIABLE
(FULL SAMPLE, N=8,281)

	1974 Linear Probability	1974 Logit	1977 Linear Probability	1977 Logit
INTERCEPT	0.771475 (7.9983)	2.05041986 (4.1425)	0.555798 (6.6981)	1.04456966† (2.1726)
BUILT	-0.032216 (6.4033)	-0.24585349 (6.9014)	-0.030029 (6.1575)	-0.27119095 (6.6828)
OWNHERE	0.107060 (4.6629)	0.65659035 (4.9071)	0.090957 (4.1328)	0.62411790 (4.3509)
DUMP	0.283919 (9.5826)	1.30671087 (9.1444)	0.184872 (9.1239)	0.91599081 (8.7858)
RACE	-0.096937 (5.2939)	-0.49761115 (5.3094)	-0.105277 (5.9640)	-0.57976291 (5.9161)
RC	-0.093653* (1.1536)	-1.01091365† (2.4637)	-0.340843 (5.1065)	-2.25296032 (5.8060)
VAC	-0.020022 (8.4024)	-0.10478498 (8.4433)	-0.0000671285* (0.0257)	-0.00619831* (0.4243)
YPC	-0.0000245941 (2.6410)	-0.00013007 (2.7604)	0.00002956914 (3.1778)	0.00014578 (2.8931)
RBUILT	0.030260 (4.5408)	0.23759451 (5.7289)	0.030547 (4.7559)	0.27479563 (5.9464)
ROWNH	0.060621† (2.1216)	0.21540184* (1.3342)	0.052181* (1.9134)	0.21532218* (1.2450)
RDUMP	-0.021321* (0.6070)	-0.14492439* (0.8544)	0.071263 (2.8121)	0.22239562* (1.7321)
RRACE	-0.118155 (5.0453)	-0.47574968 (4.0410)	-0.029511* (1.3241)	-0.10146545* (0.8307)
R^2	.1225	.0992	.0976	.0840
F	104.92		81.27	
χ^2		1059.69		823.70
D_{yx}		.411		.375
Correct(%)		67.2		74.2

(Figures in parentheses are absolute values of t-statistics for linear probability model, asymptotic t-statistics for logit model.)
All coefficients significant at 1% level unless noted as follows:
†Significant at 5% level. *Not significant at 5% level.

TABLE 5

DICHOTOMOUS QUALITY VARIABLE, LINEAR PROBABILITY MODEL

TESTS OF SIGNIFICANCE

VARIABLES	F STATISTICS Full Sample 1974	1977
RBUILT, ROWNH, RDUMP, RRACE	12.2876	9.1537
RC, RBUILT, ROWNH, RDUMP, RRACE	40.4137	22,5184
Shift: RC(1977 - 1974)	5.6458†	

†Significant at 5% level.
All others are significant at 1% level.

the magnitude of the rent control coefficients in both samples is shown to be statistically significant at the 5% level. Therefore, both hypotheses are supported by the results.

Other variables generally perform as expected, with the exception of positive coefficients on RBUILT and ROWNHERE. In addition, RDUMP is of a different sign (but insignificant) in 1974. In 1977, vacancy rate (VAC) loses significance, while per capita income (YPC) becomes significantly positive. Among interaction variables, ROWNH and RRACE lose significance in 1977.

The F statistics of the equations of both samples of the linear probability model are sufficiently large to lead to rejection of the hypothesis of joint insignificance of the independent variables (critical value of F is 2.24 at 1% significance level). The low R^2 values are most likely due not only to the problems with microdata described above, but are also an inevitable result of the nature of a linear probability model. This is because, whatever the predicted probability that a given choice will be made, an action is either chosen (so actual probability is one) or it is not (so actual probability is zero). Thus, even if a model predicts a very high probability of a choice's being made, there will be an error due to the difference between predicted probability and the actual value of the choice variable. Donald G. Morrison (1972) has explored this problem in order to develop an upper limit for R^2 in models using dichotomous dependent variables. Morrison assumes actual probabilities follow a beta distribution, which can be made to take on a unimodal, bimodal, or uniform shape, depending on one's selection of parameters. In the case of the uniform distribution, assuming predicted probabilities equal actual probabilities (so that the model is "perfect" in that it minimizes mean square error), the upper limit for R^2 is one third.[12] As the distribution of probabilities moves away from uniformity toward a unimodal form, the upper limit decreases from one third toward zero. Since in practice there is the added problem that models are not perfect so that predicted and true probabilities differ, additional downward pressure is exerted on the value of R^2.

A difficult problem arises in attempting to determine the goodness of fit of the logit model. The usual R^2 criteria do not apply here, so it is necessary to look for an alternative measure. Although there are several possibilities,[13] McFadden's R^2, a measure of proportion of log likilihood explained, will be calculated here.[14] The formula is

$$\text{McFadden's } R^2 = 1 - \frac{\ell(\beta)}{\ell(0)} \qquad (12)$$

where $\ell(\beta)$ is the log likelihood function for the parameter estimates and $\ell(0)$ is the maximum log likelihood subject to the constraint that all coefficients are equal to zero. The values for these statistics are shown in Table 4. In addition, the model chi-squares are also shown. These are computed as follows:

$$\chi^2 = \ell(0) - \ell(\beta) \qquad (13)$$

This gives an analogous statistic to the F statistic in the linear regression model, and the values for all the logit equations run are significantly different from zero.

---

[12] See, however, Arthur S. Goldberger (1973), who disagrees with Morrison's characterization of a perfect model. In Goldberger's view, such a model would predict every outcome perfectly, thus yielding an upper bound of unity.

[13] See Amemiya (1981), pp. 1502-7.

[14] See McFadden (1974) and Amemiya (1981), p. 1505.

Rather than concentrate on goodness of fit statistics of dubious value, it may be well to look at the predictive ability of the estimated models. Of interest will be the rank correlation between predicted probability and observed value of the dichotomous variable. The statistic that will be used here is Somer's D_{yx}, a nonparametric measure of association.[15] In explaining this measure, it is first necessary to introduce the concepts of concordance and discordance. Concordant pairs refer to two observations, one of which involves a correct prediction of unsatisfactory, the other of which involves a correct prediction of satisfactory. Similarly, discordant pairs involve matched incorrect predictions. The D_{yx} statistic is computed by dividing the difference between total concordant and total discordant observations and predictions by the sum of concordant and discordant pairs plus a penalty factor for pairs tied on predicted values.

Using this statistic yields, as shown in Table 4, a D_{yx} of .411 for 1974 and .375 for 1977. The interpretation is that the difference between consistently (correctly) ordered and inconsistently (incorrectly) ordered pairs as a percentage of the total is, first, positive, and, second, greater than zero. The first point means that the model predicts in the right direction, the second that the prediction is not simply random and that there exists a tendency toward monotonicity. In other words, this statistic attempts to measure the strength of the relationship, that is, agreement, between predicted and observed values of the quality variable. A useful aspect is that it takes values from -1 to +1. However, one should not immediately conclude that squaring it will yield a statistic directly analogous to the standard linear regression R^2, which measures the tendency toward linearity

[15] See Leo A. Goodman and William H. Kruskal (1972) and G. David Garson (1971), chap. 9.

of a relationship. Rather, Somer's D_{yx} is simply a rank correlation that permits evaluation of each equation's relative predictive ability. In other words, it is more concerned with association than significance. Finally, percentage of total predictions that are correct is 67.2 in 1974 and 74.2 in 1977. By these criteria, the logit equations for both 1974 and 1977 perform better than simple goodness of fit statistics would indicate.

III. Concluding Comments

According to the evidence presented in this paper, rent controls have been associated with lower housing quality than would be the case in the absence of controls. Further, this effect has been shown to grow over time. The more important implication is that, contrary to the claims of advocates, second generation rent controls have not been successful at avoiding quality deterioration.

The association of controls with lower quality does not, however, mean that there is an unambiguous welfare loss to tenants. A complete analysis would attempt to compare benefits to tenants due to wealth transfers from landlords with costs in the form of lower quality. In order to conduct such an analysis, a usable quality measure must be developed, and it is here that future efforts should focus.

REFERENCES

Amemiya, Takeshi. "Qualitative Response Models: A Survey." Journal of Economic Literature 19 (December 1981): 1483-1536.

Blumberg, Richard E.; Robbins, Brian Quinn; and Baar, Kenneth K. "The Emergence of Second Generation Rent Controls." Clearinghouse Review (August 1974): 240-9.

Domencich, Thomas A., and McFadden, Daniel. Urban Travel Demand. Amsterdam: North-Holland, 1975.

Frankena, Mark. "Alternative Models of Rent Control." Urban Studies 12 (October 1975): 303-8.

Garson, G. David. Handbook of Political Science Methods. Boston: Holbrook Press, 1971.

Gilderbloom, John. Moderate Rent Control: The Experience of U.S. Cities. Washington, D.C.: Conference on Alternative State and Local Policies, 1980.

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

Goldberger, Arthur S. "Correlations Between Binary Outcomes and Probabilistic Predictions." Journal of the American Statistical Association 68 (March 1973): 84.

Goodman, Leo A., and Kruskal, William H. "Measures of Association for Cross Classification, IV: Simplifications of Asymptotic Variances." Journal of the American Statistical Association 67 (June 1972): 415-21.

Judge, George E.; Griffiths, William E.; Hill, R. Carter; and Lee, Tsoung-Chao. The Theory and Practice of Econometrics. New York: John Wiley and Sons, 1980.

Kearl, J. R.; Pope, Clayne L.; Whiting, Gordon C.; and Wimmer, Larry T. "A Confusion of Economists?" American Economic Review 69 (May 1979): 28-37.

Linneman, Peter. "Some Empirical Results on the Nature of the Hedonic Price Function for the Urban Housing Market." Journal of Urban Economics 8 (July 1980): 47-68.

McFadden, Daniel. "Conditional Logit Analysis of Qualitative Choice Behavior." In _Frontiers of Quantitative Economics_, pp. 105-42. Edited by P. Zarembka. New York: Academic Press, 1974.

Mengle, David L. "Rent Control and Housing Quality: An Empirical Analysis of the Effect of Second Generation Rent Controls on the Quality of Rental Housing." Ph.D. dissertation, University of California, Los Angeles, 1983.

Moorhouse, John C. "Optimal Housing Maintenance Under Rent Control." _Southern Economic Journal_ 39 (July 1972): 93-106.

Morrison, Donald G. "Upper Bounds for Correlations Between Binary Outcomes and Probabilistic Predictions." _Journal of the American Statistical Association_ 67 (March 1972): 68-70.

Rydell, C. Peter; Barnett, C. Lance; Hillestad, Carol E.; Murray, Michael P.; Neels, Kevin; and Sims, Robert H. _The Impact of Rent Control on the Los Angeles Housing Market_. N-1747-LA. Santa Monica, California: The Rand Corporation, August, 1981.

U.S. Congress. Congressional Budget Office. _Federal Housing Policy: Current Programs and Recurring Issues_, by Martin D. Levine. Background Paper. Washington, D.C.: Government Printing Office, 1978.

Wieand, Kenneth F. "The Performance of Annual Housing Survey Quality Measures in Explaining Dwelling Rentals in 20 Metropolitan Areas." _American Real Estate and Urban Economics Association Journal_ 11 (Spring 1983): 45-68.

Yildez, Orhan. "Policies for Improving Housing Quality: A Comparison of Demand and Supply Subsidies." Ph.D. dissertation, Rand Graduate Institute, 1983.