N-1744-HUD

May 1982

# SPECIFICATION BIAS IN HOUSING PRODUCTION FUNCTIONS

Kevin Neels

## HOUSING ASSISTANCE SUPPLY EXPERIMENT

A RAND NOTE

*This Note was prepared for the Office of Policy Development and Research, U.S. DEPARTMENT OF HOUSING AND URBAN DEVELOPMENT, under Contract No. H-1789. Its views and conclusions do not necessarily reflect the opinions or policies of the sponsoring agency.*

Rand
SANTA MONICA, CA. 90406

Published by The Rand Corporation

## PREFACE

This note was prepared for presentation at the twenty-seventh North American Meetings of the Regional Science Association, held in Milwaukee, Wisconsin, November 14-17, 1980. It draws on data collected and prepared by The Rand Corporation as part of the Housing Assistance Supply Experiment, sponsored and funded by the Office of Policy Development and Research, U.S. Department of Housing and Urban Development (HUD), under Contract No. H-1789. The note is a product of basic research on housing market behavior, sponsored by HUD under Grant No. H-5099RG.

The author thanks Naihua Duan and C. Peter Rydell for their comments and suggestions, Penny Post for editorial help, and Karen Stewart for typing the draft.

## SUMMARY

To produce housing, a wide variety of inputs are necessary: land, structural capital, different forms of energy, water and sewer services, management, and other factors. Yet for reasons of analytical convenience and/or lack of data, researchers investigating the production function for housing services have rarely taken a comprehensive view of the production function. Instead they have usually focused on land and capital and ignored current inputs, or used inputs that were highly aggregated, such as land and nonland. Oversimplifying specification in this way clearly leads to bias in the results, but as yet we do not know how large these errors are or how they have shaped findings about production functions.

This study examines the effects of some of the more common forms of simplification assumed in analyses of the production function for housing. To establish a standard, a translog production function was fit to data drawn from the Housing Assistance Supply Experiment, taking land, capital, energy, and other current inputs as its factors. We then considered the effects of aggregating or omitting factors of production and adopting more restrictive functional forms.

Our results indicate that measurement of the first order effects in the production function is relatively insensitive to the way inputs are aggregated or to the use of restrictive functional forms. Omitting factors of production such as energy and other current inputs reduces the function's degree of homogeneity and overstates the importance of land. The second order effects in the production function appear to be less stable. All forms of specification error have the effect of pushing the estimated elasticities of substitution closer to one.

## CONTENTS

TABLES

# I. INTRODUCTION

In the last ten years the production function for housing services has received an increasing amount of attention. In a seminal article in 1964, Muth first estimated the elasticity of substitution between land and capital in housing production. Since then a number of other studies have appeared. Investigations of housing production have been hampered by the lack of adequate data, however. Researchers have therefore found it necessary to impose severe constraints on both the scope and form of the analysis.

Koenker (1972) and Sirmans, Kau, and Lee (1979) fit the parameters of housing production functions with land and capital as their inputs. Muth used the same specification,* both in his original paper (1964) and in a later investigation based upon more extensive data (1971).

Rydell published one of the very few papers that explicitly mentioned the role played by current inputs to production (1976). His model included three factors: land, capital, and "services," a composite factor including energy, water, sewer service, garbage collection, janitorial inputs, and management expenses. The form he used, however, was a three-factor Constant Elasticity of Substitution (CES) production function. The CES representation restricted elasticities of substitution between all pairs of inputs to be constant and equal. We don't know how this restriction affected his results.

The only effort made so far to relax the assumption of constant and equal elasticities of substitution in the three-factor context was reported in another paper by Muth (1973). He attempted to fit the parameters of a nested CES function in which land and capital were first

*Researchers have sometimes justified production function specifications that consider only land and capital by claiming that they are analyzing the production function for "real estate," a durable good that is then combined with current inputs to produce a flow of housing services. Many writers, however, never even consider the question of whether the output of their production function is a stock or a flow. Roger Koenker (1972) asserts that there is effectively no difference between the two. In this study, we consider only flows.

combined to produce a composite good called "real estate" that was in turn combined with current inputs to produce housing services. He defined current inputs very broadly to include expenditures for energy, water and sewer service, repairs, taxes, and insurance. Although the study was ambitious in its conception, it was seriously flawed in its execution. The highly aggregated set of data, drawn partly from new and partly from existing housing, did not allow him to generate precise estimates of the parameters of the production function.

Neither Rydell nor Muth paid any attention to the question of whether all current inputs are related to the production process in the same way and belong together in a single factor of production. Particularly in view of Muth's extremely broad definition of current inputs (1973), such aggregation seems highly questionable.

Although more information about housing has become available in recent years, it seems likely that investigators will be working with meager data for some time to come. It is also reasonable to assume that restrictive functional forms will continue to be used because they are more manageable, both theoretically and empirically. There is nothing intrinsically wrong with using models that only approximate conditions in the real world. However, the analyst needs to be aware of how close his approximation is and how it is affecting his results. It is currently difficult to make such judgments.

This study attempts to provide more information by using an exceptionally rich set of data on the nature of the housing production process. To establish a standard of comparison, we estimate parameters for a translog production function for housing services which takes as its arguments land, capital, energy, and other current inputs. The study then goes on to investigate the effect on estimates of output and substitution elasticities of adopting successively more restrictive specifications for the production function.

The results suggest that the CES function does not adequately represent the relationships that characterize the production of housing services. Current inputs to production have to be taken into account; although energy and other current inputs behave differently, according to our results, and therefore deserve separate treatment, we also found that the error introduced by combining them is small compared to that caused by leaving them out altogether.

The next two sections lay out the structure of the general and the restricted models respectively. The fourth section discusses our data sources and assumptions in developing measures for inputs and output. The fifth section discusses estimation procedures for the various models; the sixth section discusses results. The findings are summarized in the conclusion.

## II. THE GENERAL MODEL

Our general model for the production of housing services specifies that the quantity of housing produced is a function of the levels of four inputs: capital, land, energy, and other current inputs.*

Capital's contribution to production is measured by the flow of services from physical improvements to the property, such as the structure and related mechanical systems, walks, driveways, garages, wells, and septic tanks. Land simply provides a site for the structure. Energy inputs supply the power for space and water heating, lighting, and the operation of appliances. Other current inputs include water and sewer service as well as janitorial and management inputs. This list is exhaustive, including <u>all</u> inputs to the production process. We believe that the general model describes the technology of production more completely than the previous work in this area.

The general model is based upon the translog functional form introduced by Christensen, Jorgensen, and Lau in 1971. It takes the form:

$$\begin{aligned} \ln Q_h = {} & B_O + B_L \ln L + B_K \ln K + B_E \ln E + B_S \ln S + B_{LL} (\ln L)^2/2 \\ & + B_{LK} \ln L \ln K + B_{LE} \ln L \ln E + B_{LS} \ln L \ln S + B_{KK} (\ln K)^2/2 \\ & + B_{KE} \ln K \ln E + B_{KS} \ln K \ln S + B_{EE} (\ln E)^2/2 \\ & + B_{ES} \ln E \ln S + B_{SS} (\ln S)^2/2 \end{aligned} \tag{1}$$

---

*The properties and derivation of the general model are described in more detail in Neels (1981).

where $Q_h$ = the quantity of housing produced,
$L$ = land input levels,
$K$ = capital inputs levels,
$E$ = energy input levels, and
$S$ = services (other current inputs).

This function can be regarded as a Taylor series approximation in logs to an arbitrary production function. It imposes no constraints on the values that can be taken by the elasticities of substitution between factors of production. Values can vary freely from one pair of inputs to another, and for a particular pair, from one mix of inputs to another. The translog thus provides an extremely flexible representation of the technology of production.

## III. RESTRICTED MODELS

Two types of a priori restrictions on the production function were investigated. One concerned functional form; the other had to do with definition of the inputs.

The form used most often in housing production research is the CES function,* which has the property of restricting all pairwise Allen partial elasticities of substitution to be constant and equal.** The well-known Cobb-Douglas production function is a special case of the CES in which all elasticities of substitution are equal to 1. The Cobb-Douglas is thus even more restrictive than the CES.

To determine the effect of highly restrictive functional forms on estimates of production function parameters, we estimated a set of models with the same notation given above but using the Cobb-Douglas function:

$$\ln Q_h = A_0 + A_L \ln L + A_K \ln K + A_E \ln E + A_S \ln S \qquad (2)$$

The Cobb-Douglas was chosen over the CES because it was easier to estimate and because the Cobb-Douglas and the translog have similar structures. In their linear forms both take the log of output as the dependent variable. In addition, the translog reduces directly to the Cobb-Douglas when the coefficients of its higher order terms equal zero. Since the CES represents an intermediate form, its results were expected to fall between those obtained by using the translog and Cobb-Douglas forms.

The second set of restrictions investigated had to do with the definitions of the factors. Because most studies of the production function for housing have commonly ignored the role of current

*This representation of the production function was used in studies by Koenker (1972) and Rydell (1976).

**Equality of all pairwise elasticities of substitution was proved by Uzawa (1962).

inputs,* we estimated parameters for a class of models whose inputs were limited to land and capital. Those studies that do acknowledge current inputs have treated them as a third composite factor.** To see what effect that aggregation has had on parameter estimates, another class of models using three factors of production (land, capital, and current inputs) was defined and estimated.

Altogether we considered three general specifications for the production function (2-factor, 3-factor, and 4-factor) and two functional forms (Cobb-Douglas and translog). These led to six specifications for which parameters were estimated.

---

*Production functions fit in this way include those by Muth (1964, 1971), Koenker (1972), and Sirmans, Kau, and Lee (1979).

**These include studies by Muth (1973) and Rydell (1976).

## IV. MEASURING OUTPUT AND INPUTS

The data on which this analysis is based were drawn from the surveys administered as part of the Housing Assistance Supply Experiment,* a large-scale social experiment designed to test the effect of an open-enrollment housing allowance program on a local housing market. Such a program was put into operation in two north central sites: Brown County, Wisconsin (whose central city is Green Bay); and St. Joseph County, Indiana (whose central city is South Bend). To measure the market response, a set of surveys was administered annually to the owners and occupants of a stratified random sample of residential properties in each site.

The data used here came primarily from the first wave of surveys of rental properties, covering the calendar year before the allowance program began.** Public records yielded information on taxes, assessments, lot size, square feet covered by the building(s), and other basic physical characteristics. Surveyed landlords provided data on rent receipts and the expenses associated with the operation of their properties. Tenants, surveyed separately, estimated their own expenses and described the interiors of their units. In addition, trained fieldworkers rated the external conditions of residential buildings in a third survey.

The unit of observation was a rental property. Those sampled were made more homogeneous by excluding residential properties with mobile homes, rooming houses, farms, and commercial units, as well as those receiving government subsidies. What remained was a random sample of regular rental housing in Brown and St. Joseph counties. Output and all inputs were expressed on a per dwelling basis.

The derivation of the analysis sample is shown in Table 1. In each of the two sites the owners of approximately 2,000 properties were

*For an overview of the Supply Experiment, see the *Fourth Annual Report of the Housing Assistance Supply Experiment* (1978).

**For Brown County, the data cover calendar year 1973; for St. Joseph County, they cover 1974.

Table 1

DERIVATION OF THE ANALYSIS SAMPLE

| | Brown County | St. Joseph County | Total |
|---|---|---|---|
| Initial sample | 1945 | 1983 | 3928 |
| Less homeowner properties | 625 | 569 | 1194 |
| Rental property sample | 1320 | 1414 | 2734 |
| Less nonregular properties | 133 | 161 | 294 |
| Regular property sample | 1187 | 1253 | 2440 |
| Less cases lacking data | 372 | 811 | 1183 |
| Analysis sample | 815 | 442 | 1257 |

SOURCE: Tabulations by the author from HASE baseline surveys of landlords, tenants, and residential buildings.

interviewed. Roughly one fourth of these were homeowners and were thus excluded from this analysis. Exclusion of properties with farms, mobile homes, commercial units, rooming houses, and subsidized units reduced the sample by another ten percent. About half of the remaining data were unusable due to incomplete information. That problem was more severe in St. Joseph than in Brown County, primarily because of the difficulty of assembling the information needed to convert energy expenditures into estimates of the physical quantities of energy consumed.

Output

The output of the production process was defined as housing services, a homogeneous good bought and sold on the rental housing market.* A measure of the quantity of housing services provided by a particular rental property was derived from its rent roll in two steps. First we

*This definition of output draws upon a long tradition in housing research. See Olsen (1969) for a discussion of the concept of housing services.

established a measure of rent that was fully comparable across all properties; then we adjusted it to correct for differences in the prices for housing services.

The initial rent measure was cash rent, the amount a tenant pays his landlord in exchange for occupancy of a rental unit. To make the measure fully comparable, first an imputation for vacancy losses was added to rent receipts. That sum was further corrected for the presence of units occupied rent-free, usually by the owners themselves but sometimes by employees who took part of their compensation in the form of a place to live. Another correction was made for differences in what the cash rent included, because cash rents were lower where tenants paid certain bills themselves.*

Most direct tenant payments were for utilities; others covered repairs made by tenants to their units. Tenants supplied estimates of both types of bills. Another type of direct payment occurred when tenants supplied their own major household appliances; an imputation was added to account for those services.

The resulting measure of rent was fully comparable across all properties. Its accuracy as a measure of the quantity of housing services produced depended upon how constant the price of services was across properties. Random differences in price would affect the precision of the analysis but not its validity, while systematic differences could produce misleading results.

Adjustments to rent were made for three distinct types of price differentials: location rents, market condition, and landlord pricing behavior. All rents were adjusted to what they would have been had the property been located in a high-quality neighborhood in the center of Green Bay without a resident landlord and with an average moveout rate. The correction factors were derived from a hedonic index of rent described in the Appendix.

*This was the same adjustment to cash rent that was made in hedonic index studies by The Urban Institute and The Rand Corporation. See Follain and Malpezzi (1980), and Barnett (1979).

Capital

Like housing services, capital services represent a rather abstract concept describing the contributions of diverse physical attributes. We have defined the contributions of capital to the production of housing as the flow of services from the structure and other improvements to the property. In measuring this flow, the logical starting point was the value of the property. However, this measure suffered from a number of defects.

The first problem was that property value refers to the market value of the whole package of land plus improvements. To treat land and capital separately in the production function, we had to divide property value between land and improvements.

The second problem was that conceptually, capital value is a poor measure of current service flows. The value of the improvements to a property is equal to the discounted sum of the future stream of net income associated with those improvements. Their value may be high because the flow of capital services is very large, because that flow is likely to continue for a long time, or because the expenses associated with the improvements are very low. It was necessary to be able to distinguish among these effects.

The third problem was that the market value of an existing residential property was not always precisely known. Rental properties change owners relatively infrequently. Recent sales prices were available in only a minority of cases. A number of studies have shown that an alternative measure--the owner estimate of property value--is unbiased on average, but very imprecise.*

These difficulties were overcome in a three-step process.** First, total property value was divided into the value of land and the value of improvements. Second, a theoretical model was constructed to account for the effects of maintenance, taxes, insurance, and

*See Kish and Lansing (1954); Kain and Quigley (1972).

**The details of this process are described in Neels and Rydell (1981).

deterioration on the value of improvements and to allow the calculation of current capital service flows. To further refine the capital services measure, we constructed an index that related capital service flows to observable physical characteristics.* The value predicted by the index served as the measure of capital's contribution to the production of housing services.**

Energy

Reported utility bills served as the basis for measuring energy inputs. Separate estimates were obtained for each energy source (electricity, gas, fuel oil, and coal) from both tenants and landlords.

Dollar expenditures for energy were converted to estimates of physical quantities by using data on energy prices and utility rates in the two sites. Average annual energy prices were computed, and average monthly expenditures were used to arrive at estimates of average monthly energy use.

All forms of energy were reduced to a common unit of measure by means of equivalence factors based upon the heat contents of the various fuels and how efficiently they could be used. When electricity is used as a source of power, virtually all of its energy is usable. When fuel oil or natural gas is burned, however, about 30 percent of the total heat content is lost through the venting of exhaust gases. Therefore all energy forms were first expressed in comparable physical units

*The magnitude of the errors-in-variables problem is discussed in Neels and Rydell (1981).

**An added advantage of the index was that it provided a more portable measure of capital inputs than could be derived from capital value estimates. Market conditions in the two counties providing data for this study were very different. Compared to Brown County, the St. Joseph County housing market was severely depressed. Bleak prospects for the future had lowered capital values in a way that would have been difficult to control for. Fitting the index in Brown County and then using it to predict capital service flows from detailed housing attributes in St. Joseph County made it easier to assure comparability.

(BTUs); then fuel oil and gas quantities were discounted by 30 percent before they were combined with electricity into a single measure of energy.*

Other Current Inputs

The quantity of non-energy current inputs was measured by the amount spent on them, after adjusting for price differences between the two counties.** Table 2 shows the expense components.

Janitorial expenses included cleaning, trash collection and removal, and yard and grounds maintenance. Management expenses, the largest component, covered the services of lawyers, accountants, and rental agents hired by the landlord as well as the cost of business phones and office supplies. Allowance was made for time the landlord spent selecting tenants, arranging for repairs, paying bills, and performing other managerial tasks.

Cash utility expenses include the cost of piped-in water and publicly provided sewer service. Noncash utility expenses were imputed for properties located in jurisdictions where these services are provided and paid for through property taxes. Because there is often little connection between the size of the tax bill and the services provided, however, property taxes were considered as a fixed levy against the income of the property rather than as an input to production. Imputations for tax-supported water and sewer services were included to maintain comparability. Since properties with their own wells and/or septic systems use a combination of their own land and capital to produce these services, their utility expenses were zero. Like Rydell

*This aggregation procedure is the same as that followed by Baughman and Joskow (1976). They used a discount factor of .5, but found that their results were insensitive to values between .2 and .7. The value of .3 used here fits these data better, and has also been used in an analysis of energy demand in Brown and St. Joseph counties. See Neels (1981).

**The price indexes used in this adjustment were obtained from Noland (1977a and 1977b).

Table 2

NON-ENERGY CURRENT INPUTS: BROWN COUNTY, 1973, AND ST. JOSEPH COUNTY, 1974

| Expense Item | Annual Expenses Per Unit[a] |
|---|---|
| Janitorial | 78 |
| Management | 138 |
| Cash utilities | 45 |
| Noncash utilities | 8 |
| Total | 269 |

SOURCE: Tabulations by the author from all HASE baseline survey data.

[a]Expressed in 1973 Brown County dollars.

(1976), we assumed that repairs and maintenance affected output only through their influence on the quantity of capital. Since capital inputs were measured directly, including repairs and maintenance among current inputs would have constituted double counting. Therefore they were excluded.

Total Current Inputs

In the four-factor function energy was considered separately from other current inputs. In that instance a very detailed analysis of utility rates made it possible to measure energy in physical units.

In the three-factor function the quantity of energy was measured by its cost, a much cruder gauge chosen to maintain comparability with other studies of current inputs to housing production. Expenditures for both current inputs and energy were adjusted for price differences between the two counties and then added together, forming a single composite factor.

Land

Land was the easiest of the four inputs to measure. Because the terrain was quite uniform in both counties, the only characteristic of a lot we needed to know was its size. Public tax records from the two counties provided those data.

## V. ESTIMATION METHODS

We differ from earlier studies in that we have estimated the coefficients of the production function directly rather than through the factor demand equations. In his initial paper on the production function for housing, Muth (1964) developed a formal model that allowed him to infer the characteristics of the production function from the observed demand for factors of production. Most later investigators, following Muth's methodology very closely, likewise estimated the production function indirectly through the factor demand equations.*

A major problem inherent in this approach, identified by Clapp in a recent paper (1979), is that it relies heavily on precise measurements of factor prices. Clapp points out that all previous studies have used land price data derived from judgments made by appraisers or tax assessors. Because land prices are inherently difficult to measure (particularly for developed parcels of land), analyses dependent upon them are subject to serious problems of errors-in-variables. Clapp argues that this has introduced substantial bias.

A second problem with indirect estimation of the production function lies in the assumptions it requires about producer behavior. Muth's original analysis postulated that all housing producers had adjusted fully to current factor prices. While that assumption may be plausible for newly constructed housing, its validity for housing built some time in the past is questionable. The inflexibility of existing housing and the high cost of adjustments to factor ratios make it extremely unlikely that the owner of an existing residential property will remain perfectly tuned to changing factor prices. It is doubtful that such departures from profit maximization will be totally random.

As an alternative, Clapp's paper goes on to propose an "engineering" approach to the production function for housing that takes the quantity

*This method has been used by Koenker (1972), Rydell (1976), and Sirmans, Kau, and Lee (1979), among others.

of output as a dependent variable. The independent variables are transformations of input levels. The regression equation is the production function itself.

Direct estimation of the production function can be subject to simultaneous equation bias, as Hoch pointed out much earlier (1958). However, Fuss, McFadden, and Mundlak (1978) have examined this problem in detail and concluded that direct estimates are subject to bias only when the error term contains a stable component that is recognized and taken into account by producers in maximizing profits. Such a component can be distinguished from purely random ones (such as weather) that affect output in ways unknown at the time that decisions about input levels are made.

The special characteristics of the process of producing housing services from existing residential structures make it unlikely that simultaneous equation bias will arise. The difficulty of altering factor ratios once a structure has been built means that existing properties will be dominated by decisions made at the time of their construction. For most of the properties examined in this study, that was a great many years ago. Most of them changed hands several times; only rarely is a property still managed by the person or firm responsible for the original decisions about input ratios. Under these conditions, the simultaneity argument loses much of its force. We have instead a recursive system in which current input levels are related to their prices in years past.

The approach adopted in the current study followed the recommendations of Clapp. All the production functions discussed below were estimated directly using ordinary least squares.

## VI. RESULTS

Regression results for six models of housing services production are shown in Table 3. Coefficients are paired with their standard deviations, which appear immediately below. The percentage of variation explained by each of the regressions appears at the bottom of the table. Because the dependent variable was in all cases the log of output per unit, the figures for $R^2$ are directly comparable.

In all cases, a substantial loss in explanatory power occurred in moving from the translog to the comparable Cobb-Douglas function. The differences between the two forms were highly significant statistically. Table 4 presents a set of F statistics for the hypothesis that the coefficients of the cross-product terms in the translog functions were equal to zero. The statistics for the two-, three-, and four-factor functions all exceeded the critical values at the .01 significance level by a comfortable margin. Clearly, the higher order forms play an important part in the translog regressions. In comparison with the translog, the Cobb-Douglas function does a poorer job of explaining variations in output.

The four-factor functions outperform the comparable three-factor functions. The difference is quite small for the Cobb-Douglas. For the translog, however, it is substantial. Here, moving from the three- to the four-factor function increases the percentage of explained variance by three points.

However, the differences between the three- and four-factor functions pale in comparison with differences between either of them and the functions taking only land and capital as inputs. The explanatory power of the two-factor functions is far lower. It appears to be quite important that current inputs be included in the production function. In precisely what form they are included is less crucial.

The significance of current inputs can be demonstrated statistically. A test of the null hypothesis for the four-factor translog function that the coefficients for all terms involving either

Table 3

COMPARISON OF PRODUCTION FUNCTION REGRESSION RESULTS

| Term | 4-fac trnlg | 4-fac Cb-Dg | 3-fac trnlg | 3-fac Cb-Dg | 2-fac trnlg | 2-fac Cb-Dg |
|---|---|---|---|---|---|---|
| K | -1.21 | .26 | -1.32 | .26 | -.77 | .26 |
| | .28 | .00 | .29 | .01 | .29 | .01 |
| E | .26 | .13 | --- | --- | --- | --- |
| | .21 | .01 | | | | |
| S | -.25 | .04 | --- | --- | --- | --- |
| | .15 | .01 | | | | |
| L | .52 | .06 | .25 | .03 | .11 | .06 |
| | .19 | .01 | .11 | .01 | .11 | .01 |
| KK | .07 | --- | .07 | --- | .04 | --- |
| | .01 | | .02 | | .02 | |
| KE | -.04 | --- | --- | --- | --- | --- |
| | .02 | | | | | |
| KS | .04 | --- | --- | --- | --- | --- |
| | .01 | | | | | |
| KL | .04 | --- | .03 | --- | .05 | --- |
| | .02 | | .01 | | .01 | |
| EE | .03 | --- | --- | --- | --- | --- |
| | .00 | | | | | |
| ES | .02 | --- | --- | --- | --- | --- |
| | .01 | | | | | |
| EL | -.01 | --- | --- | --- | --- | --- |
| | .01 | | | | | |
| SS | .01 | --- | --- | --- | --- | --- |
| | .01 | | | | | |
| SL | -.02 | --- | --- | --- | --- | --- |
| | .01 | | | | | |
| LL | -.04 | --- | -.01 | --- | -.02 | --- |
| | .01 | | .00 | | .00 | |
| C | --- | --- | -.63 | .20 | --- | --- |
| | | | .28 | .01 | | |
| KC | --- | --- | .05 | --- | --- | --- |
| | | | .03 | | | |
| CC | --- | --- | .06 | --- | --- | --- |
| | | | .01 | | | |
| CL | --- | --- | -.03 | --- | --- | --- |
| | | | .01 | | | |
| Const. | 8.43 | 4.50 | 11.13 | 3.91 | 8.29 | 5.01 |
| | 1.59 | .13 | 1.71 | .13 | 1.31 | .13 |
| RSQR | .43 | .37 | .40 | .36 | .27 | .23 |

SOURCE: Regression analysis by the author of the HASE baseline survey data.

NOTE: K = Log of capital
E = Log of energy
S = Log of non-energy current inputs
L = Log of land
C = Log of total current inputs

Figures following coefficient estimates are standard errors.

Table 4

SIGNIFICANCE OF THE HIGHER ORDER TRANSLOG TERMS

| Number of Factors | F Statistic | Degrees of Freedom | Critical Value[a] |
|---|---|---|---|
| 4 | 12.91 | 10,1242 | 2.32 |
| 3 | 14.56 | 6,1247 | 2.80 |
| 2 | 23.73 | 3,1251 | 3.78 |

SOURCE: Computed by the author from the regression results reported in Table 3.

[a]Significant at the .01 level.

energy or other current inputs are collectively equal to zero yields an F statistic of 50.22 with 9 and 1242 degrees of freedom. The critical value at the .01 level is 2.41, indicating that the null hypothesis can be soundly rejected. A similar test for the terms of the three-factor translog involving current inputs yields an F statistic of 92.50 with 4 and 1247 degrees of freedom. The critical value at the .01 level in the latter case is 3.32. These results indicate very strongly that the two-factor functions are seriously misspecified.

Although the percentage of the total variation in output explained by the models provides a way of identifying the types of misspecifications substantially affecting the production function, it offers no information about how the qualitative characteristics of the estimated functions have been altered. In order to compare the qualitative characteristics of the different models, it is necessary to transform the coefficients into more readily interpretable measures.

The economic theory of production emphasizes the importance of the partial derivatives of output with respect to the inputs as descriptors of a production technology. The exact values of these partial derivatives (or, as they are more often called, marginal products) are sensitive to the units of measurement. A less arbitrary measure is provided by the output elasticity of a factor, which is defined as the percent change in output associated with a 1 percent change in the level of the input. We can therefore rank the various factors of production in order of importance according to the values of their output elasticities.

In the Cobb-Douglas function, the output elasticity is equal to the coefficient for the log of that factor. In the translog, the output elasticity is a linear combination of the logs of the inputs. In particular, the output elasticity for capital is given by:

$$e_{hk} = B_K + B_{KK}\ln K + B_{KE}\ln E + B_{KS}\ln S + B_{KL}\ln L \qquad (3)$$

where $e_{hk}$ is the output elasticity for capital. The remaining terms of the formula come from Eq. (1). The formulas for the other output elasticities are defined analogously.

Estimates of the output elasticities for the six models are shown in Table 5. In the three translog models, they have been computed at the sample means.

The predominance of capital in all six models conforms with both popular perceptions of the nature of the production process and the results of other analyses of housing production. The translog functions tend to assign a higher output elasticity to capital than does the Cobb-Douglas function. However, in general, output elasticities for capital were not very sensitive to the definition of the factors.

The general model assigned the second most important role to energy, as did the four-factor Cobb-Douglas. The importance of energy in explaining output is reflected also in the importance to the three-factor models of the composite factor of all current inputs. Although all models agree in emphasizing the significance of energy inputs, these inputs are treated somewhat differently by the four-factor translog and Cobb-Douglas models. The four-factor translog assigns an output elasticity to energy that is significantly higher than that estimated by the Cobb-Douglas. In this case, using a highly restrictive functional form appears to have a harmful effect.

The various models differ greatly in their treatment of land. The estimate of land's output elasticity provided by the general model, .03, was quite low. The restricted models assigned a more significant role to land, although they disagreed quite a bit about the exact figure. Estimates varied by more than a factor of three. Apparently estimates of the marginal product of land are quite sensitive to the way in which the production model is specified.

The elasticity of substitution between two inputs is defined as the proportional change in the ratio of their input levels divided by the proportional change in the ratio of their marginal products. Intuitively, this quantity can be viewed as a measure of how flexible the

Table 5

OUTPUT ELASTICITY ESTIMATES

| Factor | 4-fac trnlg | 4-fac Cb-Dg | 3-fac trnlg | 3-fac Cb-Dg | 2-fac trnlg | 2-fac Cb-Dg |
|---|---|---|---|---|---|---|
| Capital | .30 | .26 | .29 | .26 | .28 | .26 |
| Energy | .20 | .13 | --- | --- | --- | --- |
| Other Current | .05 | .04 | --- | --- | --- | --- |
| Total Current | --- | --- | .22 | .20 | --- | --- |
| Land | .03 | .06 | .06 | .03 | .10 | .06 |
| Sum | .58 | .49 | .57 | .49 | .38 | .32 |

SOURCE: Computed by the author from the regression coefficients presented in Table 3.

production technology is regarding the ways in which the two inputs can be combined. Much of the work done to date on the production function for housing has been directed toward the problem of measuring the elasticity of substitution between land and capital. When there are more than two factors for a production function, the elasticity of substitution between a pair of inputs can be defined in a number of different ways, depending upon the assumptions made about the other inputs.* If the levels of output and all other inputs are held constant, the resulting measure is the direct elasticity of substitution. Values for the direct elasticity of substitution for all pairs of inputs and all models are shown in Table 6. For the translog functions, the elasticities of substitution were computed at the sample means.

For the general model, the elasticities of substitution fall into three distinct categories. First, all those involving land were estimated to be well below 1. Elasticities of substitution for capital versus non-energy current inputs and energy versus non-energy current inputs were estimated to be much higher, and somewhat above 1 in value. Finally, the estimate of the elasticity of substitution between capital and energy was above 2 and well beyond any of the other values in the table.

The Cobb-Douglas appears to provide a very poor representation of the technology of housing production. Its form constrains the elasticities of substitution between all pairs of inputs to be equal to 1. The results for the general model indicate that this is a severe restriction. The estimated elasticities of substitution vary from one pair of inputs to another, and several are quite different from 1.

The favored form for research into housing production has been the CES function. It is more general than the Cobb-Douglas in that it allows the elasticity of substitution to take values other than zero. However, like the Cobb-Douglas it requires that elasticities of substitution between all pairs of inputs be equal. Both the three-factor and four-factor translog functions indicate that the latter restriction is inconsistent with the data. Elasticities of substitution

*The three most prominent measures are defined in McFadden (1963).

Table 6

ESTIMATES OF THE DIRECT ELASTICITIES OF SUBSTITUTION
BETWEEN FACTORS OF PRODUCTION

| Factor Pair | 4-fact trnslg | 4-fact Cb-Dgl | 3-fact trnslg | 3-fact Cb-Dgl | 2-fact trnslg | 2-fact Cb-Dgl |
|---|---|---|---|---|---|---|
| K vs E | 2.14 | 1.00 | -- | -- | -- | -- |
| K vs S | 1.27 | 1.00 | -- | -- | -- | -- |
| K vs L | .32 | 1.00 | .67 | 1.00 | .66 | 1.00 |
| E vs S | 1.37 | 1.00 | -- | -- | -- | -- |
| E vs L | .36 | 1.00 | -- | -- | -- | -- |
| S vs L | .58 | 1.00 | -- | -- | -- | -- |
| K vs C | -- | -- | 1.49 | 1.00 | -- | -- |
| C vs L | -- | -- | .95 | 1.00 | -- | -- |

SOURCE: Computed by the author from the regression coefficients presented in Table 3.

NOTE: K = Log of capital
E = Log of energy
S = Log of non-energy current inputs
L = Log of land
C = Log of total current inputs.

between land and capital are below 1. Those between current inputs and capital are well above 1. It is impossible to describe these relationships accurately using a CES function.

The only estimate in Table 6 comparable to previous results is that for the elasticity of substitution between land and capital. In an early paper (1964), Muth estimated that parameter to be equal to .75. Later, using much more extensive data, he revised it downward to .50 (1971). A year after that, Koenker published an estimate of .71 based on micro data (1972). Rydell arrived at an estimate of .50 (1976). Most recently, Sirmans, Kau, and Lee have published an estimate of .83 (1979).

The results of previous research are generally consistent with the estimates shown in Table 6 in that all except those derived from the Cobb-Douglas function are well below 1. Interestingly, the estimate provided by the general model was well below the range covered by the published results, while the estimates provided by the two misspecified translog models fell neatly in the middle. It is possible that the previously published studies may have been strongly influenced by their failure to represent the production function adequately.

The estimate for the elasticity of substitution between land and current inputs provided by the three-factor translog function fell between those provided by the general model for capital versus energy and capital versus non-energy current inputs. The three-factor model did not give this intermediate result, however, for the substitution elasticities involving current inputs and land. The estimate provided here by the three-factor model was much higher than either of the two estimates provided by the general model.

In sum, the effects of misspecification on estimates of the elasticities of substitution have been biased in all cases toward 1.

Measuring the statistical significance of the differences between the various models is difficult. The formulas for the elasticities of substitution between the factors of production are highly nonlinear in the regression parameters, making it difficult to compute their standard errors. However, the precision of the substitution elasticities implied by the general model has been previously analyzed (Neels, 1981a) using a simulation procedure. The estimates for the three substitution elasticities involving land were quite precise, with standard errors ranging from .15 to .30. The substitution elasticities among capital, energy, and services were much fuzzier in comparison, with standard errors ranging from .6 to 2.0.* On the basis of that evidence it appears that the differences between the translog models in their descriptions of the role played by land are in some sense significant. In the case of the other substitution elasticities the case is less clear. For capital and energy the differences are strongly suggestive, but not conclusive. The others have probably not been measured precisely enough to warrant firm conclusions.

*The exact figures were: capital/energy, .57; capital/services, .89; and energy/services, 2.27.

## VII. CONCLUSIONS

This study shows that misspecification of the production function for housing leads to bias in estimates of its parameters. More important, the study has identified the nature of the bias which accompanies the types of misspecification most common in the literature on housing production.

It appears that when the marginal products of the different factors are of primary interest, the Cobb-Douglas function may be useful. The output elasticities assigned to the various factors by the Cobb-Douglas and translog functions were generally similar. In both functions, capital emerged as most important, followed closely by energy; land and other current inputs played only relatively minor roles. Though numerical estimates of the output elasticities were not the same in the two functions, they were close.

The Cobb-Douglas does a poor job, however, of representing the second-order characteristics of the production function for housing. Estimates of the substitution elasticities obtained from the translog were far from unity, the value to which they are constrained by the Cobb-Douglas. In analyses where the values taken by the substitution elasticities are important, the Cobb-Douglas function is likely to produce seriously misleading results.

In its treatment of substitution elasticities, the CES function is little better. Estimates obtained from the translog for the elasticities of substitution for capital vs. land and capital vs. energy differed from each other by far more than either differed from unity. The fact that the CES function constrains all elasticities of substitution to be equal appears therefore to be quite a serious limitation.

The results obtained from the general model support the idea that energy and other current inputs deserve separate treatment. The two factors appear to behave quite differently. The estimated elasticity of substitution between capital and energy was much higher than that between capital and other current inputs. However, the damage done by combining these two inputs fades into insignificance compared to that which results from excluding them altogether.

We can infer from these results that the benefit of differentiating among energy inputs and moving to a five-factor production function is likely to be small.* Although a more detailed description of inputs would undoubtedly clarify still further the nature of the production process, it would not greatly increase the explanatory power of the regression or produce qualitatively different results.

The effect of omitting or combining factors of production was to bias the sum of the output elasticities toward zero, and the point estimates for the substitution elasticities toward 1.

The production function measures physical relationships, which means that the results would apply to housing in other times and places where general building practices are similar. Specifically, they are most relevant to small rental properties located in the northern part of the country. How appropriate they would be for housing located in southern climates where the mix of inputs is very different, or in areas where the rental stock is dominated by large structures, would have to be tested separately.

---

*Since electricity and fossil fuels are often used for very different functions, an a priori case could be made for treating them as separate factors of production.

Appendix

## MEASURING HOUSING SERVICE PRICES AND LOCATION RENTS

Variation in housing prices within the production function analysis sample was measured using a hedonic index of rent.* The general specification for the rent index was:

$$\ln R = a_0 + a_1X_1 + \ldots + a_nX_n + b_1P_1 + \ldots + b_mP_m \qquad (A.1)$$

where R was equal to gross rent per unit per year. The $X_i$ variables described attributes of the rental units in the sample. They included measures such as the number of rooms per unit and the overall condition of the structure. Where the attribute measures had a well-defined cardinal scale they were generally entered in logarithmic form.

It is customary in hedonic index studies to interpret the coefficients of the attribute variables as measures of the implicit market prices for those attributes. To develop measures of overall price variation it was necessary to view them somewhat differently, however. It was assumed that the expression:

$$a_0 + a_1X_1 + \ldots + a_nX_n$$

provided a rough measure of the quantity of housing being produced. Using this measure to control for quantity made it possible to observe variation in rents due to price differentials.

The $P_i$ variables measured dimensions of housing service price variation. They fell into three general groups. The first captured

*There have been a great many studies of rent based upon hedonic index methodology. See Barnett (1979) for an example of this literature.

differences in the overall price levels of the markets included in the sample. The second group measured location rents within these markets. The third accounted for differences in landlord pricing behavior.

The data for the rent analysis were taken from the HASE surveys described in the text. The unit of observation was the residential property, and the dependent variable was the log of gross rent for the property (described in the text) divided by the number of units on the property. Property level attribute variables were obtained by averaging across all residential units. Parcels with farms, rooming houses, mobile homes, or commercial units were omitted from the sample.

The sample used to estimate the coefficients of equation (A.1) contained 1510 observations. The explanatory variables were able to account for slightly over half of the total variation in the dependent variable. The overall F statistic for the equation was 57.2.

The attribute variables and their coefficients are shown in Table A.1. Floorspace (measured in square feet) was obtained from public records. The appliance stock measure was defined as the purchase price of the major appliances present in 1976 dollars. It was not entered in logarithmic form because there were a few units in the sample without any major appliances. Structural quality, rated on a continuous scale ranging from one up to four for the best quality, was constructed by averaging ten separate ratings for walls, windows, roofs, sidewalks, floors, ceilings, and foundations. Seven were provided by trained fieldworkers as part of a survey of residential buildings. Three came from the occupants of the units on the property. Because the structural quality variable had no definable units, it was entered directly rather than in logarithmic form.

Lotsize was measured in square feet and was truncated at a value of one half acre per unit. Values above this amount were assumed to have no effect upon rents.

Table A.2 shows the estimated coefficients for the variables related to location rents and landlord pricing behavior. The meanings of most of these are fairly clear. The vacant land measure refers to vacant lots within the urbanized area; in St. Joseph County, they are

Table A.1

RENT INDEX REGRESSION COEFFICIENTS FOR
ATTRIBUTE VARIABLES

| Variable | Coefficient | T-Statistic |
|---|---|---|
| Log of rooms per unit | .3654 | 16.39 |
| Log of one plus bathrooms per unit | .3473 | 6.87 |
| Log of floorspace per room | .0993 | 6.49 |
| Appliance stocks | .00015 | 7.72 |
| Structural quality | .0525 | 3.09 |
| Presence of wood or composition siding | -.0623 | -5.20 |
| Presence of central heat | .0539 | 3.36 |
| Presence of a garage | .0220 | 1.93 |
| Presence of a thermostat | .0875 | 3.96 |
| Presence of a lobby | .0697 | 3.11 |
| Log of lotsize per unit | .0320 | 2.96 |

SOURCE: Regression analysis by the author of HASE baseline survey data.

Table A.2

RENT INDEX REGRESSION COEFFICIENTS FOR PRICE DISEQUILIBRIUM AND LOCATION VARIABLES

| Variable | Coefficient | T-Statistic |
|---|---|---|
| Price disequilibrium: | | |
| Presence of a resident landlord | -.0526 | -3.13 |
| Moveouts per unit | .0736 | 7.25 |
| Site I location variables: | | |
| Distance from the central business district | -.0117 | -4.22 |
| Overall neighborhood quality | .1036 | 5.07 |
| Presence of farms nearby | .0392 | 1.21 |
| Site II location variables: | | |
| Distance from the central business district | -.0104 | -4.45 |
| Overall neighborhood quality | .0360 | 1.00 |
| Presence of consumer shops on blockface | -.0229 | -1.34 |
| Presence of mixed commercial residential land on blockface | -.0316 | -1.34 |
| Presence of vacant land on blockface | -.0402 | -2.33 |
| Presence of farms nearby | -.1322 | -2.77 |

SOURCE: Regression analysis by the author of HASE survey data.

often the end result of abandonment and demolition. The farm variable measures farms in the immediate vicinity of the property. The neighborhood quality measure is an average of separate ratings of the condition of other residential buildings, yards and landscaping, and overall cleanliness. On a scale of one to four, it was measured by the size of the HASE neighborhood, an area roughly 3 to 4 times the size of a census tract. All location variables were transformed by subtracting from them their site-specific means.

Table A.3 shows the estimated coefficients for the submarket dummy variables, along with the implied price indexes. These indexes measure nominal prices. The St. Joseph County coefficients are generally higher than the Brown County coefficients because the former measure 1974 rents while the latter measure 1973 rents. Between 1973 and 1974 consumer prices in general went up 11 percent.

For the production function analysis, three price indexes were needed. The first measured market-related price differentials. For this, the indexes in the last column of Table A.3 were used. The second and third measured location rents and variations in the prices charged by landlords respectively. These were obtained by using the variables and coefficients shown in Table A.2. The location rent index had values of one at the site-specific mean values of the location variables; higher or lower values resulted in higher or lower values respectively for the index. The index measuring differences in landlord pricing behavior took values of one for nonresident landlord properties with average moveouts in the year covered by the data, more than one with more moveouts, and lower with fewer moveouts or with a resident landlord.

Table A.3

RENT INDEX COEFFICIENTS AND PRICE INDEXES
FOR HOUSING MARKETS

| Market | Dummy Variable Coefficient | T-Statistic | Price Index |
|---|---|---|---|
| Brown County: | | | |
| Single family | [a] | [a] | 1.000 |
| Multiple family | -.0197 | -1.04 | .981 |
| Central South Bend: | | | |
| Single family | .1168 | 5.26 | 1.124 |
| Multiple family | -.0367 | -1.31 | .964 |
| Rest of St. Joseph County: | | | |
| Single family | .1652 | 8.12 | 1.180 |
| Multiple family | .0329 | 1.24 | 1.033 |

SOURCE: Regression analysis by the author of HASE survey data.

[a]Not available. This was the excluded category.

# BIBLIOGRAPHY


Barnett, C. Lance, Using Hedonic Indexes to Measure Housing Quantity, The Rand Corporation, R-2450-HUD, October 1979.

Baughman, Martin L., and Paul L. Joskow, "Energy Consumption and Fuel Choice by Residential and Commercial Customers in the United States," Energy Systems and Policy, Vol. 1, No. 4, 1976.

Christensen, Laurits, Dale Jorgensen, and Lawrence Lau, "Conjugate Duality and the Transcendental Logarithmic Functions," Econometrica, Vol. 39, No. 4, 1971 (abstract).

Clapp, John M., "The Substitution of Land for Other Inputs," Journal of Urban Economics, Vol. 6, No. 1, 1979.

Follain, James R., and Stephen Malpezzi, Dissecting Housing Value and Rent: Estimates of Hedonic Indices for Thirty-Nine Large SMSA's The Urban Institute, URI-28700, Washington, D.C., February, 1980.

Fourth Annual Report of the Housing Assistance Supply Experiment, The Rand Corporation, R-2302-HUD, May 1978.

Fuss, M., Daniel McFadden, and Yair Mundlak, "A Survey of Functional Forms in the Economic Analysis of Production," in M. Fuss and Daniel McFadden (eds.), Production Economics: A Dual Approach to Theory and Applications, Vol. I, North Holland, Amsterdam, 1978.

Hoch, Irving, "Simultaneous Equation Bias in the Context of the Cobb-Douglas Production Function," Econometrica, Vol. 26, No. 4, 1958.

Kain, John F., and John N. Quigley, "A Note on Owner's Estimate of Housing Value," Journal of the American Statistical Association, Vol. 67, No. 340, 1972.

Kish, Leslie, and John B. Lansing, "Response Errors in Estimating the Value of Homes," Journal of the American Statistical Association, Vol. 49, No. 267, 1954.

Koenker, Roger, "An Empirical Note on the Elasticity of Substitution Between Capital and Land in a Monocentric Housing Market," Journal of Regional Science, Vol. 12, No. 2, 1972.

McFadden, Daniel, "Constant Elasticity of Substitution Production Functions," Review of Economic Studies, Vol. 30 (2), No. 83, 1963.

Muth, Richard, "The Derived Demand Curve for a Productive Factor and the Industry Supply Curve," Oxford Economic Papers, Vol. 16, No. 2, 1964.

Muth, Richard, "The Derived Demand for Urban Residential Land," Urban Studies, Vol. 8, No. 2, 1971.

Muth, Richard, "Capital and Current Expenditures in the Production of Housing," in C. L. Harriss (ed.), Government Spending and Land Values, The University of Wisconsin Press, Madison, Wisconsin, 1973.

Neels, Kevin, The Derived Demand for Energy in the Production of Housing Services, Ph.D. thesis, Department of City and Regional Planning, Cornell University, Ithaca, N.Y., 1981.

Neels, Kevin, Families, Houses, and the Demand for Energy, The Rand Corporation, N-1542-HUD, February 1981.

Neels, Kevin, and C. Peter Rydell, Measuring Capital's Contribution to Housing Services Production, The Rand Corporation, P-6587, February 1981.

Noland, Charles W., Indexing the Cost of Producing Housing Services: Site I, 1973, The Rand Corporation, N-1088-HUD, April 1981 (first issued as WN-9022-HUD, January 1977).

Noland, Charles W., Indexing the Cost of Producing Housing Services: Site II, 1974, The Rand Corporation, N-1118-HUD, April 1981 (first issued as WN-9736-HUD, May 1977).

Olsen, Edgar O., "A Competitive Theory of the Housing Market," American Economic Review, Vol. 59, No. 4, 1969.

Rydell, C. Peter, "Measuring the Supply Response to Housing Allowances," Papers of the Regional Science Association, Vol. 37, 1976.

Sirmans, C. F., James B. Kau, and Cheng F. Lee, "The Elasticity of Substitution in Urban Housing Production: A VES Approach," Journal of Urban Economics, Vol. 6, No. 1, 1979.

Uzawa, Hirofumi, "Production Functions with Constant Elasticities of Substitution," Review of Economic Studies, Vol. 29 (4), No. 81, 1962.

RAND/N-1744-HUD