MEASURING THE COST OF SHELTER FOR HOMEOWNERS: THEORETICAL AND EMPIRICAL CONSIDERATIONS

Robert Gillingham*

I. Introduction

RECENT economic developments have aroused substantial interest in the treatment in the Consumer Price Index (CPI) of the cost of shelter for homeowners.[1] From December 1977, when the latest version of the CPI was introduced, until December 1980, the all-items CPI increased at an average annual rate of 11.6%, while the homeownership component increased at an average annual rate of 16.2%. Relative to all of the other goods and services in the CPI, the homeownership component has increased by 17.5% over the same time period. If the relative price of homeownership had remained constant the growth rate of the CPI would have been reduced to 10.1%. The question which has been raised is whether the rapid relative increase in the homeownership component, which has had such an important impact on the CPI, truly reflects changes in the cost of shelter. This question is not only important, but difficult, encompassing many subsidiary questions and auxiliary issues. The purposes of this paper are threefold: (1) to outline briefly a conceptual framework for the CPI, which leads to a straightforward specification of what the shelter component of the CPI should measure, (2) to evaluate the theoretical properties of alternative procedures designed to approximate this measurement objective and (3) present empirical evidence on the operational difficulties involved in pursuing a new approach to shelter cost measurement. Two main conclusions are reached. First, on both theoretical and empirical grounds, a "rental equivalence" approach to measuring shelter costs for owner-occupants is preferred. Second, an estimated rental equivalence measure has grown more slowly over (at least) the past six years, than the official CPI homeownership component. Given the way in which the CPI is used to escalate both private and public expenditures, these results demonstrate that the choice of measurement technique has important distributional implications.

II. A Conceptual Framework for the CPI

The conceptual framework which will underlie our analysis of shelter costs is based on two basic assumptions. The first assumption is that the appropriate theoretical construct for measuring consumer prices is the cost-of-living index (see Pollak (1971) and Fisher and Shell (1972)). Within this framework, the objective of the CPI is to measure changes, for a given consumer, in the cost of achieving a given level of satisfaction. For the purposes of this paper, the important aspect of this assumption is that it specifies the cost of living as a function of the determinants of the consumer's level of satisfaction or, in other words, the arguments of the consumer's utility function. The second important assumption is that the consumer's welfare is determined by the flows of consumption services which he receives, where the services can be either (1) directly provided, (2) obtained coincidentally with the consumption of a nondurable good, in which case the distinction between good and service is unnecessary, or (3) obtained from the use of a durable good owned by the consumer. In each case, satisfaction is derived from the act of consumption; ownership of a source of consumption services (a durable good) produces no additional satisfaction. The purchase of a durable good is an "investment," designed to provide consumption services over a future time span.

Within this framework, the CPI can approximate a cost-of-living index by measuring the cost

Received for publication October 29, 1981. Revision accepted for publication August 10, 1982.

*Bureau of Labor Statistics.

The author would like to thank Steven Cobb, Kenneth Dalton, John Greenlees, Walter Lane, W. John Layng, Janet Norwood, Robert Pollak, Jack Triplett and anonymous referees for helpful discussions and/or comments on earlier drafts. He would also like to thank Walter Lane for help in preparing the experimental user cost indexes and Herbert Cover and Eugene Luckritz for help in preparing the rent data base. The views expressed are those of the author and do not reflect the policies of the BLS or the views of other BLS staff members.

[1] See, for instance, Blinder (1981), Council of Economic Advisers (1981) and Gordon (1981). This topic has also been the subject of numerous editorials and several congressional hearings. After this paper was written, the BLS announced that the rental equivalence approach to measuring shelter costs for homeowners recommended in this paper will be phased into the official CPI over the next three years.

over time of the market basket of services *consumed* in the base period. For the services provided by directly purchased services and nondurable goods this implies observing market prices and transaction levels in the base period, as well as the subsequent time path of market prices. For the services provided by durable goods owned by consumers, the implicit price of the services must be estimated, since market transactions do not take place each time the service is consumed.

The remainder of this paper will analyze this estimation problem for the case of shelter services provided by owner-occupied homes, comparing two alternative approaches—user cost and rental equivalence—which have received substantial support as solutions to this problem. I will start by defining user cost in the simplest case—in a certain, taxless world with perfect markets—and proceed to outline the conceptual and empirical complications which arise when these assumptions are dropped.

III. Theoretical Framework for Measuring the Cost of Shelter

In a world with active rental and resale markets, no uncertainty, and no friction costs, the user cost of a house in a given period is

$$C(t) = r(t) \cdot P(t) - A(t) + Z(t) \tag{1}$$

where r is the (single) rate of interest in period t, P is the average price of the house in period t, A is equal to the change in the average price over the period and Z represents all other cost components. In other words, the user cost is defined as the opportunity cost of holding and using the house, $r \cdot P + Z$, less the increase in its value. In equilibrium the rental price of the house, R, will be equal to the user cost, and, since frictions have been assumed away, the rent received by a landlord will equal the rent paid by a tenant. Thus, in a perfect world

$$R^L(t) = C(t) = R^T(t) \tag{2}$$

where the superscripts L and T denote landlord and tenant, respectively.

Under the conditions I have assumed, measurement of the value of the flow of shelter services from a house is trivial. It can be measured with information from either rental or resale and money markets, and it does not matter whether the information refers to buyers' or sellers' prices. The problems arise when one attempts to measure the cost of shelter for homeowners in a more complicated setting, when the exact form of the user cost function is more difficult to define and the equalities defined above need not hold.

To lay out this problem more clearly, I will drop the assumption of perfect certainty, thereby allowing for a structure of differential asset yields. I will also relax the assumption of no friction costs to allow for the possibility that the net rent received by a landlord may be less than the rent paid by a tenant, the difference representing, for instance, the cost of advertising. I do assume that there is a sufficiently active rental market that there is some price at which each homeowner can rent shelter services equivalent to those provided by his own home and some strictly positive price at which another consumer would be willing to rent his house. Under these conditions, the user cost measure can be redefined as

$$\begin{aligned} C(t) = {} & r_e(t)E(t) + r_m(t)M(t) \\ & - A(t) + Z(t) \end{aligned} \tag{3}$$

where M and E are mortgage and equity amounts which sum to P, r_m is the mortgage interest rate, and r_e is the opportunity cost of equity capital.

The relationship between user cost defined in this manner and the alternative rent measures defined above is now ambiguous, and depends critically on the manner in which the opportunity cost of equity capital is defined.[2] Certainly R^T must be greater than or equal to R^L, but depending on the manner in which one chooses to define and estimate r_e, the relationship between each of the rent measures and C is uncertain.

The variables included in the redefined user cost function are all conceptually and operationally straightforward with one crucial exception—r_e. Unfortunately, estimates of user cost are also sensitive to alternative definitions of this variable. Several somewhat "natural" alternatives for defining r_e have been suggested in Gillingham (1973a), McFadyen and Hobart (1978), Muth (1972, 1975),

[2] It is important to note that the appropriate measurement objective within the CPI or cost-of-living framework is actual, total user cost. Of equal interest would be a measure of expected, marginal user cost, which could be used in estimating models of consumer behavior. In this case, measures of expected capital gains and expected rates of return on equity would be appropriate components. The differences between the appropriate methods to compute and use these two measures are confused by Dougherty and Van Order (1981 and 1982).

and Steiner (1961). In Gillingham (1973a), it was suggested that r_e be estimated as an ex post internal rate of return defined by the identity

$$R^L(t) + A(t) \equiv r_e(t)E(t) + r_m(t)M(t) + Z(t), \qquad (4)$$

where R^L is an estimate of the market rental which an owner could receive for his house. Alternatively, one might argue that the appropriate internal rate of return be defined by substituting R^T for R^L in equation (4).[3] In either case the resulting estimate of the user cost, which we will call C^*, reduces to an implicit rent, and the following relationship holds:

$$R^L(t) \le C^*(t) \le R^T(t). \qquad (5)$$

McFadyen and Hobart (1978), Muth (1975), and Steiner (1961) suggest that alternative rates of return which consumers either receive or pay—such as consumer debt, savings account, mortgage interest and bond rates—be used to construct the user cost function. With this approach, depending upon the particular rates of return included, the resulting user cost estimate need not be bracketed by the two rent variables.

The problem of selecting an appropriate definition for r_e reduces to a fundamental question concerning the interpretation of the user cost of housing services. Use of an ex post internal rate of return excludes returns to investment in housing assets from the shelter services index except as these returns affect rent levels. To give an example, assume a change in government policy which spurs an increased rate of appreciation in house prices without affecting rent levels, interest rates or the other cost components included in Z. In this case, both actual appreciation and the ex post internal rate of return on housing would increase, but the user cost of housing services, as measured with the internal rate, would be unaffected. Use of an alternative rate would result in an unrealistically low estimate of the user cost of services, since the increase in the appreciation rate would be counted as an offset to shelter costs when, in fact, no change had occurred in the price charged for shelter services in rental markets. The use of an alternative rate of return in the user cost function implies that, even with a given potential rent for the services of a house, the user cost of those services is dependent on the specific financial and operating costs and price trend of that house. I feel it is more reasonable to allow the rental market to determine the value of the housing services and allow variations in other user cost factors which are specific to that house to be offset by variations in the internal rate of return. The relationship between an alternative rate and an internal rate will provide information about the profitability of housing investment (both in a particular house and housing in general), but the use of an alternative rate in the user cost function will not yield a measure of *service* cost which is necessarily consistent with rental market information.

To put this argument another way, it is plausible to contend that a user cost measure is a conceptually viable estimate of the value of the flow of shelter services only if it is bracketed by R^L and R^T. It cannot be less than R^L since a homeowner always forgoes this amount when he lives in his own house, and it cannot be greater than R^T since a homeowner always has the alternative of obtaining equivalent housing services at this price. For these reasons, given a sufficiently active rental market, a user cost measure is viable only if it is bounded by R^L and R^T, and this condition will obtain in general only if an appropriately defined internal rate is used to measure the opportunity cost of equity.[4]

Accepting the existence of active rental markets, the foregoing discussion leads to the conclusion that, from a theoretical point of view, rental market information is a necessary input—either direct or indirect—into the construction of a user cost measure of shelter costs. Thus, a rental equivalence approach provides a simpler, more direct measure. Before concluding that the rental equivalence ap-

[3] The "landlord" and "tenant" rents, as well as the user cost measure, discussed in this section are all measured gross of any special tax treatment. Simple adjustments can be made to the user cost measure suggested here to obtain a user cost measure net of tax which incorporates the fact that the imputed rents of owner-occupants are not taxed and landlords can deduct certain expenses which an owner-occupant cannot.

[4] It is important to note that the arguments in this section do not depend on the existence of equilibrium in either housing rental or asset markets. Furthermore, the ex post internal rate of return measure which we seek is an accounting measure and bears no necessary relationship, in the short run, to the equilibrium or expected rates of return on housing investment. The approach suggested here does not require knowledge of the long-run interrelationship among the expected rate of return, expected capital gains and expected user cost.

proach is also better from an operational point of view, however, I will analyze two empirical questions: (1) Do alternative estimates of the opportunity cost of equity to homeowners yield, in practice, reasonable approximations to a user cost index? (2) Can a reasonable rental equivalence measure, which takes into account the differences between the stock of owner-occupied and renter-occupied houses, be constructed from rental market data? Both of these questions, and especially the second, are complex and difficult to answer definitively. The remainder of this paper will deal with each of these questions in turn, however, providing what empirical information is currently available.

IV. Alternative Estimates of User Cost

Using the general user cost formula, defined above as

$$C(t) = r_e(t)E(t) + r_m(t)M(t) - A(t) + Z(t), \qquad (6)$$

I experimented with alternative estimates of the various components—especially r_e, r_m and A—in an attempt to develop a user cost function which would provide an estimate of the trend in shelter cost which is reasonably consistent with rental market information. (In the remainder of this section, the word "reasonable" is used only to represent this limited criterion.) The basic problem in this approach is that two of the components in the above formula—r_e and A—are historically volatile and, *ceteris paribus*, correlated. Although the measurement of house price levels, and thus appreciation, is difficult, it is possible to construct reasonable estimates of current appreciation which do, in fact, accurately reflect the historical volatility of this series.[5] Without using information from rental or housing investment markets, however, it is not possible to capture the presumably correlated variation in the internal rate of return on housing. In other words, it is impossible to estimate a user cost measure which exhibits reasonable short-term movements when current appreciation rates are included in the measure.

As a result of this empirical anomaly, my experiments were focused on developing a user cost measure which would provide a reasonable estimate of the trend movement in user cost without exhibiting the unrealistic short-term fluctuations which characterize a user cost measure which includes current appreciation rates. To do this I experimented with (1) alternative averages of appreciation rates to estimate the trend movement in appreciation, and (2) alternative averages of mortgage interest rate indexes to estimate the trend movement in both the opportunity cost of equity capital and mortgage costs. It might be hoped that a user cost measure which incorporates these trend measures for both appreciation and the opportunity cost of equity would provide a more reasonable trend estimate for the user cost of housing. Furthermore, an index constructed in this fashion could be constrained to reduce unreasonable short-term volatility, characteristic of several of the alternative measures considered, which would cause severe problems in both the use and interpretation of the index.

Tables 1 and 2 summarize, very briefly, the basic findings of the analysis.[6] These tables are based on user cost simulations which incorporated five alternative estimates for the appreciation rate and four alternative specifications for both r_e and r_m. The simulations use CPI component indexes to measure Z. In the first table, in which the current mortgage interest rate is used for both r_e and r_m and the CPI house price series is used for $E + M$,

[5] It is worth reemphasizing that the objective of this paper is to estimate actual, historical user cost and not expected future cost (cf. footnotes 2 and 4). There is, presumably, a correct formulation of equation (6) incorporating expectations of r_e and A to estimate the latter, but this measure would not necessarily be an adequate estimate of the former. In measuring the current, historical user cost of owner-occupied housing one might question (1) whether current or historical mortgage interest rates are the appropriate rate with which to measure mortgage costs and (2) how leverage should be treated. The framework developed in this paper makes it clear that these issues are not important. Since the services of a house have an (implicit) market value, this value, along with whatever mortgage rate and leverage percentage are chosen, will determine the appropriate return on equity as a residual. Thus, for example, higher mortgage interest rates result in lower equity returns, *ceteris paribus*, and vice versa. The choice of mortgage interest rate or leverage percentage can be governed by matters of convenience and whether one wants to distinguish equity returns that stem from financing differences from other equity returns.

[6] The indexes in tables 1 and 2 update similar tables presented in Gillingham (1980). They differ slightly because the average appreciation rate for a year is the average over all twelve months rather than the average of the last month of each quarter. The house price series was constructed by using the CPI house price index to adjust the price of a "reference" house, which was valued at $14,000 in March of 1964. The appreciation rates were then computed from this series.

TABLE 1.—ESTIMATED USER COST INDEXES

Period	Appreciation Rate Averaged Over Current Period	1 year	3 years	5 years	10 years
December 1964	100.0	100.0	100.0	100.0	100.0
December 1965	57.9	97.5	97.0	100.7	101.2
December 1966	53.1	116.2	113.6	112.4	113.5
December 1967	39.3	102.8	111.4	109.7	117.4
December 1968	−13.5	89.3	114.8	119.7	129.1
December 1969	49.4	66.5	105.2	123.0	140.3
December 1970	62.1	58.1	89.1	120.0	148.7
December 1971	130.7	122.8	79.5	100.2	134.9
December 1972	155.9	149.4	108.0	102.0	135.0
December 1973	54.5	168.2	165.3	130.4	158.2
December 1974	−131.3	57.4	157.4	148.2	173.6
December 1975	152.1	59.6	109.5	147.0	163.5
December 1976	150.8	190.0	96.3	142.8	153.8
December 1977	53.6	107.9	126.9	132.4	156.7
December 1978	−20.6	50.8	146.8	119.4	172.0
December 1979	16.1	−20.6	96.5	151.7	217.1
December 1980	522.0	176.0	131.2	219.9	291.4

Note: Current mortgage interest rates used for both r_e and r_m.

TABLE 2.—ESTIMATED USER COST INDEXES

Period	Interest Rate Averaged Over Current Period	1 year	3 years	5 years	10 years
December 1964	100.0	100.0	100.0	100.0	100.0
December 1965	101.2	100.7	100.1	99.6	101.3
December 1966	113.5	108.2	104.1	102.4	104.8
December 1967	117.4	116.9	111.7	108.4	110.4
December 1968	129.1	123.9	117.5	111.9	112.5
December 1969	140.3	135.3	123.4	115.4	112.8
December 1970	148.7	147.1	132.7	121.7	113.5
December 1971	134.9	136.5	139.9	128.9	117.9
December 1972	135.0	134.6	140.5	134.2	122.0
December 1973	158.2	143.9	139.5	140.4	127.9
December 1974	173.6	161.5	143.1	140.9	128.1
December 1975	163.5	159.8	148.7	137.5	127.8
December 1976	153.8	158.7	158.0	143.6	134.1
December 1977	156.7	153.8	156.1	148.4	137.4
December 1978	172.0	158.9	149.4	146.9	132.8
December 1979	217.1	189.4	156.2	147.0	129.2
December 1980	291.4	269.4	205.3	175.2	145.3

Note: All indexes incorporate a 10-year unweighted average of appreciation rates.

the impact of alternative estimates of appreciation are displayed. For the index in column 1, appreciation is estimated by applying current appreciation rates to current (constant quality) house prices. For the indexes in columns 2 through 5 the appreciation rate is approximated by one-, three-, five- and ten-year unweighted averages of lagged appreciation rates. Comparison of these indexes demonstrates the extreme impact of appreciation on the user cost measure. Even with the five-year average the index exhibits extreme and unlikely dips, which are only partially damped when appreciation is averaged over ten years. Given historical appreciation patterns, longer averaging periods not only smooth the index but also result in a substantial increase in the estimated change in user cost from 1964 to 1980. Unfortunately, there is no theoretical basis on which to choose from among the alternatives.

In table 2, which incorporates the 10-year weighted average appreciation rate for all indexes, the current mortgage interest rate and four alternative weighted averages of mortgage interest rates are used to represent r_e and r_m. As with alternative estimates of appreciation, the choice of the interest rate to represent r_e and r_m has a substantial impact on the index, though longer averaging periods for interest rates make the index rise more slowly. Again, without recourse to rental market information the choice of r_e is essentially arbitrary. In summary, the indexes in tables 1 and 2 demonstrate the difficulty inherent in developing a theoretically defensible user cost index without information on implicit rents.[7] None of the indexes track rent movements, as measured by the CPI rent index, closely. Unless rents for the types of units which are typically owner-occupied move quite differently than the rents of typical renter units, the methods summarized in tables 1 and 2 do not appear promising. The question of whether

[7] Dougherty and Van Order (1982) estimate two user cost functions similar to those presented in tables 1 and 2, using the current mortgage interest rate for r_e in both and estimating the rates of appreciation as two alternative moving averages of the rates of change in an index of new, single-family house prices. There are many problems with their analysis—e.g., their theoretical model is essentially the oversimplified model in equation (1) and they focus on marginal cost, which is inappropriate for the CPI, rather than average cost. Furthermore, although they imply that their approach is a viable alternative for the CPI, they have ignored the inequality constraint suggested in section II of this paper and have used essentially arbitrary smoothing procedures similar in effect to those reflected in table 1. Although they do not report user cost indexes in Dougherty and Van Order (1982), they do report similarly specified user cost indexes in Dougherty and Van Order (1981). These indexes fluctuate widely and increase at an unrealistically low rate in comparison to rents. Furthermore, as in table 1, the indexes are very sensitive to the choice of procedure for estimating appreciation rates. DeLeeuw (1981, p. 135) describes the Dougherty and Van Order (1981) procedures for estimating user cost indexes as "so volatile that they make even the total CPI rate of change look like a random walk at times."

rent movements are homogeneous across different types of units will be addressed in the next section.

V. Estimation of a Rental Equivalence Index

The results described in the previous section give added support for a rental equivalence approach to measuring shelter costs for homeowners. In this section I will describe attempts to estimate a rental equivalence index using data from the ongoing residential rent survey conducted by the Bureau of Labor Statistics (BLS). Much of the discussion concerning the feasibility of constructing a rental-equivalence measure has centered around the question of whether data from rental units can be used to develop measures which are representative of owner-occupied units. One extreme possibility is that the implicit rent of owner-occupied units changes at the same rate as the actual rents of rental units and that little, if any, additional effort is required to obtain a usable rental equivalence index. The BLS currently publishes an experimental CPI (CPI-U, X1) which uses the CPI rent component without modification as a proxy for a shelter cost index for homeowners. To the extent that the rate of change in rents does not vary across tenure classes, the BLS experimental index should be a reasonable rental equivalence index.[8] The BLS is quick to point out, however, that there is no basis for accepting this assumption and that additional research is necessary to analyze variation in rent movement across housing unit characteristics (directly) and tenure class (indirectly). The results reported below are a step in this direction.

A. *Empirical Specification and Description of Data*

The indexes presented below are based on the assumption that the rent that is (or could be) charged for a housing unit is a function of the characteristics of (1) the unit itself, (2) the neighborhood in which it is located, (3) the tenants, and (4) the lease, which specifies the services included in rent. Under this assumption I estimated the relationship between these characteristics and the rent of the unit.[9] I interpret the function specified above as the outcome of the interaction of supply and demand schedules for dwelling units with different mixes of these characteristics. As such, it is best interpreted as a tool for estimating the rent of a housing unit with a particular set of characteristics which, though unobserved (or observed in insufficient quantity), is actually traded in the rental market. It does not provide an adequate basis for estimating the rent that a particular type of unit, which does not exist or is not traded, would have if it did exist and was traded.[10] Thus, my basic assumption is that the distributions of rental- and owner-occupied units by characteristic, though characterized by different relative frequencies, cover essentially the same area in characteristics space.

Data on rent and the characteristics of the unit, the tenant and the lease were drawn from the BLS rent survey. In this survey, housing units are visited every six months, and rent and characteristics information are collected for both the current and previous months. Approximately one sixth of the rental units (a "panel") are visited each month. Neighborhood characteristic data for metropolitan areas, as of 1970, were obtained by merging the BLS rent sample with Census tract information.[11] The dependent variable in each regression was the logarithm of contract rent. In this specification, a given absolute change in the level of a quality attribute has a constant percentage impact on contract rent. Use of this form can be interpreted as (1) an assumption that the conditional variance of rent is proportional to its expected value and (2)

[8] I abstract here from other problems in the rent index. Cf. the discussion of "aging" bias below. It is the BLS rent index which is used to deflate both the "rental equivalence" and actual rent aggregates in the Personal Consumption Expenditures component of the National Income and Product Accounts.

[9] I used this approach because it is quite easy to estimate the regression model and construct indexes. I am not, however, suggesting that this is a viable, much less the best, approach to computing a rent index for the CPI on a monthly basis. The CPI rent survey design allows the same units to be visited over time. Presumably, an operational, monthly rental equivalence index would make use of the panel nature of the survey to explicitly or implicitly reweight the rent relatives for individual units to obtain a rental equivalence index. I did not follow this approach for two reasons: (1) it would be beyond the scope of this paper to accurately simulate CPI procedures and (2) I wanted to evaluate other problems, such as the aging bias, and this evaluation requires use of a different methodology.

[10] Cf. Gillingham (1973b, chapter 2) or Pollak (1979) for additional discussion of this issue.

[11] Neighborhood characteristics are available only for metropolitan areas, so housing units in nonmetropolitan areas were dropped from the analysis. For simplicity, I also dropped housing units which had no private entrance, no kitchen, no bathroom or no heating facilities whatever, as well as housing units in Anchorage and Honolulu.

an assumption that the quality, and thus the cost, of the attribute is related to the rent level. Most of the independent variables are dichotomous. With the exception of median neighborhood income, all continuous variables are measured in level form, though quadratic terms are included for the age of the dwelling unit and number of rooms to allow for nonlinear relationships between age and size and the logarithm of rent. The logarithm of median neighborhood income was used because this variable is measured in 1969 dollars and relative differences are likely to be more meaningful over time than absolute differences. Separate regressions were estimated for each region to allow for coefficient variation across regions. Geographic dummy variables were included in each regression to allow for interarea price variation. A full set of the independent variables included in the analysis is provided in appendix table 1.

Rent survey data have been processed on a continuing basis since 1975. Each regression was run on twelve months of data, with the periods overlapping by six months so a "chain" biannual index could be constructed.[12] Eleven separate regressions were estimated for each region. Each of the regressions was of the form

$$\ln(R) = \sum_i x_i\beta_i + \sum_i x_i d\lambda_i + \sum_j z_j\alpha_j, \tag{7}$$

where R is the contract rent of the unit, x is the vector of variables with coefficients which vary between six-month subperiods, z is a vector of other variables and d is a dummy variable equal to one in the second of the six-month subperiods over which the regression is pooled.[13] The coefficient vector λ is used to construct the index between the two six-month subperiods. The results of the estimation are summarized in appendix tables 2 through 4, with the first table focusing on β and a subset of α, the second focusing on λ and the last providing summary statistics. On average, the regressions explained over 60% of the variation in the dependent variable and, to the extent that I have strong priors, the coefficients have predominantly correct signs. Perhaps the most interesting result is that the average coefficients on all of the slope interaction terms are extremely small relative to their standard deviations, and none of these terms has a coefficient which is consistently of one sign. For each regression, I tested the null hypothesis that all of these coefficients were zero. In only five out of forty-four cases can the null hypothesis be rejected at the 0.05 level. These results give little evidence of a consistent, long-run variation in the movement of rents for different types of units. Complete sets of coefficient estimates, as well as additional tables summarizing the results, are available from the author.

B. *Construction of Rent and Rental Equivalence Indexes*

The regression coefficients on the slope and intercept interaction terms form the basis for computing both the rent and rental equivalence index-

[12] In specifying and estimating the regression model, I encountered two problems. First, although the rental units were randomly assigned to collection panels, preliminary analysis indicated nonrandom variation in average rent levels across panels which was not accounted for by the regression model. This problem caused regressions run on data from short intervals (less than six months) to be somewhat erratic, so all six panels were included in each regression. Second, I faced conflicting objectives in specifying the model. On the one hand, I wanted to include a large number of independent variables to obtain greater explanatory power and avoid the exclusion of potentially important explanatory variables. On the other hand, the variations in coefficients across periods, much of which could be random (especially for variables with insignificant and unstable coefficients), determines the movement of the indexes constructed from the regression. As a compromise, in each regression I pooled two six-month periods and allowed only a small number of coefficients to vary between the two subperiods.

The coefficients which are allowed to vary, through the use of slope and intercept interaction terms, are those on the intercept and the variables which define the age and size of the unit, whether the unit is detached, and the median income of the neighborhood in which it is located, as well as a time dummy variable indicating whether the rent was paid in the first or second quarter of each six-month subperiod. With the exception of the intercept and the time variable, these variables were chosen because (1) they are consistently, statistically significant in the regression analysis and (2) they vary substantially, on average, between the owner and renter subgroups of the housing stock. The last variable is included to allow for more flexibility in the time path of rent change. The analysis was also carried out allowing all the coefficients to vary between subperiods. The indexes obtained from these regressions, though slightly more erratic than those presented in this paper, are still quite reasonable.

[13] I also experimented with additional variables but dropped several, including those relating to utilities and furnishings included in rent, that had inconclusive coefficient estimates. It would have been useful and interesting to include a variable denoting the rent control status of an apartment, but this information was unavailable. The SMSA dummy variables should account in part for the prevalence of rent control. The fact that detached unit rents, which are not subject to rent controls, did not increase more rapidly than rents on other housing unit types would indicate that rent controls, to the extent that they existed, did not exert significant downward pressure on rent changes.

es. For the regression specification used in this analysis, the index can vary with the size, age and median neighborhood income of the unit, as well as between detached and other structure types. For each region, indexes were computed for the regional mean values, for each tenure group, of these variables. The rental unit means were computed from the sample of rental units priced for the BLS Rent Survey in the first six months of 1978. The owner unit means were computed over the owner-occupied units in the 1972–73 Consumer Expenditure Survey (CES), a survey which collected information equivalent to that collected on the rent survey. The neighborhood unit information for the CES sample was obtained by matching "pseudo-tract" data—data from geographic units similar, but not identical, to Census tracts—with the CES data. The value of the variable which defines the quarter in which the rent was charged was arbitrarily set to 0.5 for all index computations. Regional mean values for the variables with slope interaction terms are presented in appendix tables 5 and 6 for renters and owners, respectively.

Estimated indexes were computed for each tenure class in each region in the following manner:

$$\ln(I_t) = \ln(I_{t-1} \cdot F_{t,t-1}) = \ln(I_{t-1}) + \sum_i c_i \lambda_i^*, \tag{8}$$

where I_t is the estimated index in time t for specification c, $F_{t,t-1}$ is the predicted geometric mean relative between periods t and $t-1$, and λ^* is the estimated coefficient vector for the pooled regression containing periods t and $t-1$. The indexes are arbitrarily set to 100 in the first half of 1978. The estimated regional indexes, from 1975 through 1980, are presented in appendix table 7, along with the official CPI rent component. (The official rent indexes are available by region only since December 1977.) Estimated national indexes are presented in table 3, along with the CPI rent and homeownership indexes. The hedonic indexes are, on the whole, quite reasonable, both at the national and regional level. While it should be reemphasized that the methodology used to construct these indexes is not a realistic procedure for producing a monthly CPI, the estimated indexes provide no reason to conclude that a rent data base, and especially one augmented to better represent owner-occupied housing, does not provide an adequate basis for producing a rental equivalence index. Contrasting the rental equivalence index presented in table 3 with those presented in tables 1 and 2 might lead one to be even more positive about the viability of a rental equivalence index.

Given the above main conclusion, there are two additional aspects of the indexes presented in table 3 and appendix table 7 which deserve emphasis. First, the hedonic rent index increases more rapidly than the official CPI rent component over the period for which both are available. Furthermore, this phenomenon persists across all regions. Although there are a number of differences between the two methodologies, a likely explanation for a substantial part of the faster increase is that the official methodology does not adjust for the fact that units in the sample are getting older and that the official index suffers from an "aging" bias. The coefficients on the level and square of age in the hedonic regressions, evaluated at the regional means of the age variable, would account for less than half of the difference between the two indexes, however, so that, to the extent that an aging bias accounts for more of the difference, it cannot be identified explicitly from the regressions.

Second, the rental equivalence index increases more slowly than the hedonic rent index over both the entire sample period and the post 1977 period. Many scenarios could be constructed to explain this phenomenon. One explanation that deserves additional investigation is that this relative movement results from (1) rents increasing less rapidly for higher quality housing units than for lower quality housing units over the recent past and (2) owner-occupied housing units being of higher quality, on average, than renter units.

These two aspects of the hedonic indexes lead to an interesting interpretation of CPI-U, X1, the experimental CPI which substitutes the CPI rent component for the homeownership index to approximate a rental equivalence index. To the extent that the official rent index is biased downward, CPI-U, X1 is also biased downward. However, to the extent that implicit rents for homeowners have increased more slowly than actual rents, this bias is mitigated. The indexes presented in table 3 would imply a rather small historical bias in CPI-U, X1, as a flow-of-services CPI, relative to that which has resulted from the use of the current treatment of homeownership in the official index.

TABLE 3.—ALTERNATIVE SHELTER INDEXES FOR TOTAL U.S. METROPOLITAN POPULATION

Period	"Official" CPI Rent[a]	Hedonic Rent	Rental Equivalence	"Official" CPI Homeownership[a]
1975 I	84.2	80.5	81.5	81.4
1975 II	86.2	83.0	84.4	84.0
1976 I	88.7	85.5	86.6	86.2
1976 II	91.0	89.3	90.7	88.3
1977 I	93.8	92.1	93.1	91.2
1977 II	96.6	96.0	97.7	95.3
1978 I	100.0	100.0	100.0	100.0
1978 II	103.5	104.4	103.7	107.0
1979 I	107.0	109.1	108.9	114.0
1979 II	111.5	114.6	112.7	125.2
1980 I	116.4	121.6	118.6	139.5
1980 II	121.6	128.2	124.6	147.6

[a] From 1978 on, the "official" BLS indexes were constructed by "subtracting" the non-metropolitan indexes from the total U.S. index, and are thus not exact, official indexes. Prior to 1978, the "official" indexes include nonmetropolitan coverage, and the rent index is based on a different rent survey.

VI. Summary and Conclusions

The purpose of this paper has been to evaluate, on both conceptual and empirical grounds, alternative methods for measuring the cost of shelter for homeowners in the Consumer Price Index. Conditional on several basic assumptions about the appropriate theoretical framework for the CPI, we concluded that the appropriate measure of housing costs is the value of the flow of shelter services, i.e., the user cost, of housing. Unfortunately, however, in a world of uncertainty and imperfect markets, user cost is an ambiguous concept, requiring assumptions about the appropriate definition of its components. Within this framework, I argued that the most useful definition of the user cost of housing requires information about the rent which a housing unit could command if it were rented and that, given this requirement, the implicit rent of an owner-occupied house could be used more efficiently to measure user cost directly.

Given this conclusion two empirical questions were then investigated: (1) Do alternative estimates of user cost, which do not make use of implicit rental information, yield empirically reasonable measures of user cost? (2) Can a reasonable rental equivalence measure, which takes into account the differences between the stock of owner-occupied and renter-occupied houses, be constructed from rental market data? The answer to the first question is a resounding "no." Under a wide set of alternative formulations, it was not possible to obtain a user cost index which was consistent with trends in rent levels. Other procedures might yield this result ex post, but there is no reason to believe that this relationship would hold in the future.

The answer to the second question is a somewhat more guarded "yes." A regression approach was used to estimate both rent and rental equivalence indexes. The procedure used yields a reasonable measure of relatively long-term—six months and over—changes in actual and implicit rent levels, but is not directly amenable to measuring accurately short-run change. With this caveat, along with the standard caution that the validity of the estimated indexes is conditional on the validity of the regression specification, I estimated both actual and implicit rent indexes. In general, the estimated actual rent index increased more rapidly than the official CPI rent component, but the estimated rental equivalence index increased more slowly than the estimated actual rent index. Though it is impossible to *prove* that a reasonable rental equivalence index can be constructed, my experiments provided no information to the contrary.

The results concerning the feasibility of a rental equivalence index are especially important when one considers the impact that such a change in procedure would have had on the CPI over the three years between December 1977 to December 1980. During that period the official CPI-U grew 38.9% while CPI-U, X1 grew "only" 32.5%. This growth in the official CPI roughly corresponds to an increase in "indexed" expenditures in the federal budget from 150.9 billion dollars in fiscal year 1979 to an estimated 231.3 billion dollars in fiscal 1982 (Council of Economic Advisers, 1981, tables 4 and 5, pp. 13–14). If, during the same period, federal expenditures had been indexed by the CPI-U, X1 or a CPI which incorporated the rent and rental equivalence indexes presented in this paper, 1982 indexed expenditures would be approximately 10 billion dollars less. It is beyond the scope of this paper to specify either the objective(s) of indexing or the appropriate index to be used. However, if a flow of services type of index is desired, the impact on indexed transfers in general, and transfers in the federal budget in particular, could be formidable indeed, and development of an operational rental equivalence index would seem to deserve a high priority.[14]

REFERENCES

Blinder, Alan, "The Consumer Price Index and the Measurement of Recent Inflation," *Brookings Papers on Economic Activity* 11 (1981).
Council of Economic Advisers, "Report on Indexing Federal Programs" (Jan. 1981).
DeLeeuw, Frank, "Discussion," in John Tuccillo and Kevin Villani (eds.), *House Prices and Inflation* (Washington, D.C.: The Urban Institute Press, 1981).
Dougherty, Ann, and Robert Van Order, "Inflation and Housing Costs," in John Tuccillo and Kevin Villani (eds.) *House Prices and Inflation* (Washington, D.C.: The Urban Institute Press, 1981).
______, "Inflation, Housing Costs and the Consumer Price Index," *American Economic Review* 72 (Mar. 1982), 154–164.
Fisher, Franklin, and Karl Shell, *The Economic Theory of Price Indexes* (New York: Academic Press, 1972).
Gillingham, Robert, "Measurement in the Consumer Price Index of the Cost of Shelter to Homeowners," Bureau of Labor Statistics, mimeograph (1973a).
______, "Place to Place Rent Measurements Using Hedonic Quality Adjustment Techniques," unpublished Ph.D. Dissertation, University of Pennsylvania (1973b).
______, "Estimating the User Cost of Owner-Occupied Housing," *Monthly Labor Review* 63 (Feb. 1980), 31–35.
Gordon, Robert, "The Consumer Price Index: Measuring Inflation and Causing It," *The Public Interest* 63 (Spring 1981), 112–134.
McFadyen, Stuart, and Robert Hobart, "An Alternative Measurement of Housing Costs and the Consumer Price Index," Bureau of Labor Statistics, *Canadian Journal of Economics* 2 (1978), 105–112.
Muth, Richard, "On the Measurement of Shelter Costs for Homeowners in the Consumer Price Index," Bureau of Labor Statistics, mimeograph, 1972.
______, "The User Cost of Owner-occupied Housing," Bureau of Labor Statistics, mimeograph, 1975.
Pollak, Robert, "The Theory of the Cost of Living Index," Bureau of Labor Statistics Working Paper no. 11 (June 1971).
______, "The Treatment of 'Quality' in the Cost of Living Index," Bureau of Labor Statistics Working Paper no. 90 (July 1979).
Steiner, Peter, "Consumer Durables in an Index of Consumer Prices," Staff Paper No. 6 in *The Price Statistics of the Federal Government* (New York: National Bureau of Economic Research ("General Series," No. 73), 1961).

[14] Several other issues, which were beyond the scope of this paper, also deserve attention in future work. The most important of these, at least from an empirical standpoint, is the treatment of taxes in a flow of services shelter index. Adequate treatment, however, requires addressing tax treatment with respect to the CPI as a whole, as well as the housing component, an approach which will be employed in a later paper.

TABLE A-1.—SPECIFICATION OF VARIABLES

A. Housing Unit Characteristics
1. AGE OF UNIT—measured in decades
2. (AGE OF UNIT)^2
3. AGE UNKNOWN—age of unit unknown
4. ATTACHED BUILDING—attached single family dwelling
5. "LARGE" BUILDING—building with 5 or more units
6. DETACHED BUILDING—single family detached building
7. NUMBER OF ROOMS—more than 9 rooms coded as 9
8. (NUMBER OF ROOMS)^2
9. TWO OR MORE BATHS
10. CENTRAL HEAT
11. CENTRAL AIR CONDITIONING
12. PARKING FACILITIES—garage or carport

B. Services and Facilities Included in Rent
1. REFRIGERATOR/RANGE—0 if neither, 1 if either, 2 if both
2. CENTRAL HEAT IN RENT
3. ROOM HEATING EQUIPMENT—room heating equipment in rent

C. Tenant Characteristics
1. NUMBER OF OCCUPANTS
2. NONBLACK OCCUPANT—head of household non-black
3. RACE UNKNOWN—race of head of household unknown

D. Neighborhood Characteristics
1. PROPORTION BLACK
2. PROPORTION SPANISH
3. PROPORTION SINGLE UNITS—Proportion of units in single family dwellings
4. PROPORTION MULTI-UNIT—Proportion of units in buildings with 5 or more units
5. PROPORTION RENTAL UNITS—Proportion of occupied units which are rented
6. PROPORTION SUBSTD PLUMBING—Proportion of units with incomplete plumbing
7. LOG (MEDIAN INCOME)—logarithm of median tract income

E. Locational Characteristics
1. RURAL—housing unit in rural area
2. CITY SIZE ('C')—housing unit in city between 75,000 and 385,000 population
3. CITY SIZE ('A')—individual city variables are used for 'A' stratum cities, lists are given in appendix table 2.

F. Other Characteristics
1. PANEL 1–PANEL 5—five variables which define the "panel" (the sixth of the sample) to which the housing unit belongs
2. QUARTER—reported rent applies to a month in either the second or fourth calendar quarter
3. TENANT REPORTER—data collected from tenant

Note: The "reference" specification is a housing unit in an elevator building in a 'B' size city (between 385,000 and 1,250,000 population and not in 'A' stratum), which also has zeroes for other dummy variables. The elevator building and 'B' stratum variables are omitted to avoid singularity.

TABLE A-2.—MEAN REGRESSION COEFFICIENTS ON HOUSING UNIT, LEASE, TENANT AND NEIGHBORHOOD CHARACTERISTICS, BY REGION

Variable	East	North Central	South	West
AGE OF UNIT	−.0549	−.0340	−.0819	−.0668
$(AGE)^2$	.0087	.0077	.0149	.0094
AGE UNKNOWN	−.3837	−.1573	−.2736	−.1897
ATTACHED BUILDING	−.0755[a]	.0166[a]	.0070[b]	−.0072[b]
"LARGE" BUILDING	.1592	.1110	.1487	.0300[b]
DETACHED BUILDING	−.0572	.0082[b]	−.0089[b]	.0350[a]
NUMBER OF ROOMS	.0631	.0936	.0960	.1153
$(\text{NUMBER OF ROOMS})^2$	−.0057	−.0091	−.0007[b]	−.0049[a]
TWO OR THREE BATHS	.2586	.1510	.1809	.1490
CENTRAL HEAT	.3579	.3637	.2202	.1098
CENTRAL AIR CONDITIONING	.1303	.1461	.0561	.0555
PARKING FACILITIES	.0621	.0687	−.0039[b]	.0252
REFRIGERATOR/RANGE	.1014	.0948	.1148	.0584
CENTRAL HEAT IN RENT	.0994	.0572	.0581[a]	.0143[b]
ROOM HEATING EQUIPMENT	.2188	.1981	.1213	.0708
NUMBER OF OCCUPANTS	.0295	.0168	.0215	.0272
NONBLACK OCCUPANTS	.0204	.0228	.1093	.0387
RACE UNKNOWN	.0303[b]	.0587	.0945	.0387
PROPORTION BLACK	−.0797[b]	−.2253[a]	−.0862[a]	−.1842[a]
PROPORTION SPANISH	.0886[b]	.2585[a]	.0763[a]	.2013[a]
PROPORTION SINGLE UNIT	−.1431[a]	.0770[a]	−.2770	−.2930
PROPORTION MULTI-UNITS	.2817	.2276	.1529	.0268[b]
PROPORTION RENTAL UNITS	−.2308[a]	.0674[b]	−.1134	.0840[a]
PROPORTION SUBSTD PLUMBING	.3365[a]	.1775[a]	−.4249	−.6492
LOG (MEDIAN INCOME)	.4115	.4312	.2792	.5036
INTERCEPT	4.4610	4.1547	4.6191	4.9275

[a] Coefficient less than twice its standard deviation across time periods.
[b] Coefficient less than its standard deviation across time periods.

TABLE A-3.—MEAN REGRESSION COEFFICIENTS FOR INTERACTION TERMS, BY REGION (STANDARD DEVIATIONS IN PARENTHESES)

Interaction Variable[a]	East	North Central	South	West
AGE OF UNIT	.0004	.0008	.0009	.0012
	(.0037)	(.0025)	(.0085)	(.0042)
$(AGE)^2$	−.0006	−.0003	−.0001	.0002
	(.0018)	(.0014)	(.0030)	(.0018)
AGE UNKNOWN	.0056	−.0001	.0017	−.0000
	(.0678)	(.0354)	(.0450)	(.0282)
DETACHED BUILDING	.0060	.0007	−.0036	−.0026
	(.0324)	(.0149)	(.0196)	(.0188)
NUMBER OF ROOMS	−.0023	−.0003	−.0006	.0003
	(.0047)	(.0045)	(.0071)	(.0042)
$(\text{NUMBER OF ROOMS})^2$	−.0006	−.0003	.0002	−.0005
	(.0023)	(.0012)	(.0029)	(.0039)
LOG (MEDIAN INCOME)	−.0023	.0008	.0017	−.0087
	(.0210)	(.0248)	(.0179)	(.0300)
QUARTER	.0026	.0022	.0026	.0026
	(.0178)	(.0146)	(.0067)	(.0079)
INTERCEPT	.0470	.0377	.0426	.0493
	(.0094)	(.0157)	(.0157)	(.0129)

[a] All variables for which slope interaction coefficients were estimated are measured as deviations from the rental unit means given in appendix table 4. Consequently, the intercept interaction coefficient gives a direct estimate of the rate of change of rents for this standard rental unit.

TABLE A-4.—SUMMARY STATISTICS, BY REGION

Statistic	East	North Central	South	West
AVERAGE R-SQUARE	.6018	.6226	.7332	.5418
AVERAGE MEAN STD ERROR	.2944	.2427	.2795	.2658
AVERAGE SAMPLE SIZE	7031	9524	6993	7326
TOTAL SAMPLE SIZE	41731	56574	41426	43615

TABLE A-5.—MEAN RENTER-OCCUPIED HOUSING UNIT SPECIFICATIONS, BY REGION

Variable	East	North Central	South	West
AGE OF UNIT	4.065	3.513	2.596	2.465
AGE UNKNOWN	0.081	0.018	0.033	0.086
DETACHED BUILDING	0.068	0.126	0.238	0.238
NUMBER OF ROOMS	4.063	4.075	4.098	3.901
LOG (MEDIAN INCOME)	9.199	9.209	9.079	9.162

TABLE A-6.—MEAN OWNER-OCCUPIED HOUSING UNIT SPECIFICATIONS, BY REGION

Variable	East	North Central	South	West
AGE OF UNIT	3.053	2.694	2.060	2.053
AGE UNKNOWN	0.052	0.035	0.056	0.030
DETACHED BUILDING	0.738	0.865	0.905	0.885
NUMBER OF ROOMS	6.364	5.935	5.916	5.875
LOG (MEDIAN INCOME)	9.316	9.311	9.128	9.327

TABLE A-7.—ALTERNATIVE RENT AND RENTAL EQUIVALENCE INDEXES, BY REGION

	East			North Central			South			West		
Period	"Official" CPI Rent Index	Hedonic Rent Index	Rental Equivalence Index	"Official" CPI Rent Index	Hedonic Rent Index	Rental Equivalence Index	"Official" CPI Rent Index	Hedonic Rent Index	Rental Equivalence Index	"Official" CPI Rent Index	Hedonic Rent Index	Rental Equivalence Index
1975 I		78.7	81.5		83.9	83.7		80.8	82.1		78.3	78.5
1975 II		83.0	87.9		85.0	84.1		83.7	82.4		80.1	82.8
1976 I		85.0	87.3		86.7	86.6		86.8	86.7		83.2	85.4
1976 II		87.5	90.0		91.8	92.4		91.2	92.1		86.6	87.9
1977 I		90.9	96.0		93.8	94.6		93.0	91.4		90.7	89.5
1977 II		94.6	98.9		97.7	99.6		97.0	95.9		94.7	96.1
1978 I	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1978 II	102.1	104.1	104.7	103.2	103.8	104.1	104.5	103.7	101.4	104.5	106.3	104.7
1979 I	104.9	107.9	107.4	106.7	109.4	110.1	107.7	107.6	108.4	109.0	111.8	110.2
1979 II	108.4	112.4	107.9	110.7	115.3	115.5	112.3	113.0	112.6	115.4	118.5	115.5
1980 I	112.6	117.4	113.5	115.2	121.4	119.9	117.2	122.3	119.3	121.7	126.1	122.9
1980 II	117.4	123.2	119.9	119.5	127.0	124.4	122.8	129.0	126.7	127.9	134.7	128.1

[a] The "official" BLS indexes were constructed by "subtracting" the nonmetropolitan indexes from the total U.S. index, and are thus not exact, official indexes.