Hedonic-Based Price Indexes for Housing: Theory, Estimation, and Index Construction

Nancy E. Wallace

Associate Professor of Real Estate, University of California at Berkeley and Visiting Scholar. I gratefully acknowledge the comments and suggestions of Brian Motley, Joe Mattey, and seminar participants at the Federal Reserve Bank of San Francisco. Any errors and omissions are my own.

Housing price indexes should not confound the effect of changes in quality with the effects of changing house prices. A recent nonparametric regression technique, loess, *allows flexible estimation of the hedonic price function and centers the estimation at fixed points, such as the beginning or ending period housing characteristics. Indexes using these estimates are consistent with the requirements of Laspeyres and Paasche price indexes. The technique is used to obtain indexes for fifteen municipalities in Alameda County from 1970:Q1 through 1995:Q1. The nonparametric hedonic-based indexes provide better controls for the effect of quality evolution on price movements than alternative methods.*

Residential real estate accounts for about 70% of the wealth portfolio of the average U.S. household. Residential housing assets also provide the collateral support for the residential mortgage market with an outstanding stock of about $3.2 trillion in 1996. In addition to the size of the housing market, the market is also notable because it is prone to boom and bust cycles. The unpredictability of these cycles introduces considerable volatility into the wealth positions of the average household and to the mark-to-market value of residential mortgages held in portfolio by financial institutions, pension funds, insurance companies, and individual investors.

In the last four or five years there has been a concerted effort to develop valuation methods that give market participants more accurate information about residential real estate price levels and returns over time. One reason for this interest is growing investor demand for measures of value and return that are comparable to the wide variety of indexes available for the bond and stock markets. A second reason is the increased sophistication of real estate investors and the more widespread use of modern tools of financial analysis, such as portfolio allocation models, option pricing models, and advances in structuring real estate investment vehicles through securitization. A final reason is the search for cost efficiencies in mortgage lending and real estate portfolio management. Cost efficiency has led to the increased use of automated appraisal and underwriting technologies and reliance on capital-at-risk models which require accurate measures of risk and return by asset class. For these reasons, many practitioners would like housing price indexes that are transaction-based and that can be produced with high levels of reporting frequency and accuracy.

Most currently available housing price indexes are transaction-based; reporting frequency and accuracy, however, remain unresolved issues. All the available strategies must contend with the fact that transactions are infrequent and that information on the terms of sale and the characteristics of the properties are costly to obtain. Choosing among existing methods to obtain housing price indexes must be done on the basis of the desired application. The choices here would include whether the index is intended to proxy the price per unit of the housing stock, whether it is intended to estimate the changing price level (or returns) of a "representative" house over time, or whether what is sought is an estimate of the value of a particular house or a portfolio of houses over time.

The purpose of this paper is to consider hedonic-based indexes of housing prices. The indexes are evaluated using a comprehensive transaction-based data set for residential sales from first quarter 1970 through first quarter 1995 for fifteen municipalities in Alameda County (171,131 transactions). The intent of this review is not to demonstrate the superiority of the hedonic-based method, but rather to highlight the empirical importance of the theoretical assumptions that underlie it. For some applications, the hedonic-based indexes would not be expected to differ greatly from strategies such as repeat sales indexes. This would be the case for applications in which there are large numbers of repeat transactions in a housing market and the market is characterized by low levels of production or remodeling. The repeat sales and hedonic methods would be expected to be equivalent if it is reasonable to assume that both the levels and prices of the underlying housing attributes, such as bathrooms and bedrooms, have remained the same over time. For other applications, however, the differences between the methods are important both theoretically and substantively. This would be the case in markets for which it is not reasonable to assume that attribute prices and levels are constant over time

The advantages of hedonic-based methods must also be evaluated relative to their cost of application. These costs vary greatly by state. States such as California have a number of high quality vendors of residential transaction data, while other states do not have these services commercially available. Thus, the appropriate choice of price index methodology also depends upon data availability.

The paper is organized into five sections. In Sections I and II, I will survey the theoretical framework for hedonic price indexes and housing price index number construction. The purpose of this overview is to highlight the assumptions required to obtain econometrically estimable price indexes and the economic theory that supports these assumptions. This conceptual framework is important because it establishes guidelines for the estimation methods and allows for meaningful interpretation of empirical results. In Section III, I will discuss two non-parametric formulations for price index composition using hedonic price functions. I apply these strategies using transaction data from Alameda County and evaluate the results. Section IV provides a graphical evaluation of the price indexes constructed from hedonic-based methods and those using repeat sales. Section V concludes.

I. HEDONIC PRICE FUNCTIONS FOR HOUSING

In economics, housing is usually treated as a heterogeneous good, defined by a set of characteristics such as square footage, bathrooms, public service amenities, and location, among many others. The number of such characteristics is indexed by j and the number of houses produced by n. The price of housing is defined by a hedonic price function, which is a mathematical relationship between the prices of the composite housing assets and the quantities of characteristics embodied in them. Thus,

$$P = h(x), \tag{1}$$

where P is an n-element vector of house prices, x is a $j \times n$ matrix of house-specific characteristics.

In the housing market, the economic decisionmaking behavior of market participants (behavior related to what is being demanded or supplied) really pertains to housing characteristics. A housing transaction is a tied sale of a set of characteristics.

To formalize the assumption that characteristics are the true arguments of the consumption- and/or production-optimization strategies of economic agents, assume for simplicity that there is only one heterogeneous good, housing, and the utility function for a household can be written as:

$$Q = Q(q(x), c), \tag{2}$$

where Q is utility, $q(\cdot)$ is a function over the housing characteristics, and c is all other homogeneous consumption goods. The production of housing assets can be represented as the joint output of a bundle of housing characteristics. Assuming the usual capital, labor, and materials (KLM) production function this can be written as:

$$t(x, K, L, M) = 0, \tag{3}$$

where $t(\cdot)$ is a transformation relationship in production.

It is well-established that the hedonic price function, $h(\cdot)$, does not represent a "reduced form" for supply and demand functions derived from the utility or production functions (Rosen 1974, Epple 1987). Instead the hedonic, $h(\cdot)$, should be thought of as the binding constraint in the optimization problems of producers and purchasers of housing.[1] Rosen (1974) shows that as long as there is increasing marginal cost of characteristics for producer/sellers and a

1. Rosen (1974) identifies special cases in which the hedonic price surface can be identified. These cases include; (1) when there is only a single type of buyer the $q(\cdot)$'s are identical so that the $h(\cdot)$ is uniquely identified by the functional form of $q(\cdot)$ and (2) when there is only a single type of seller the $t(\cdot)$'s are identical so that the $h(\cdot)$ is uniquely identified by the functional form of the $t(\cdot)$. In the former case the hedonic

constraint on unbundling the attribute package, the hedonic function is likely to be nonlinear. The nonlinearity of the hedonic constraint implies that relative characteristics prices are not fixed and instead are uniquely determined for each buyer by the buyer's location on the hedonic surface.

To illustrate the problem, consider Figure 1. It shows two nonlinear hedonic price contours for houses with two characteristics ($P_1 = h(x_1,x_2)$ and $P_2 = h(x_1,x_2)$) at a given time period. The $P_1(P_2)$ contour describes all possible types of houses that sell for price $P_1(P_2)$ and are composites of the two characteristics, x_1 and x_2 , such as square footage and number of rooms. The slope of the $P_1(P_2)$ contour defines the marginal purchase costs for the respective characteristics.

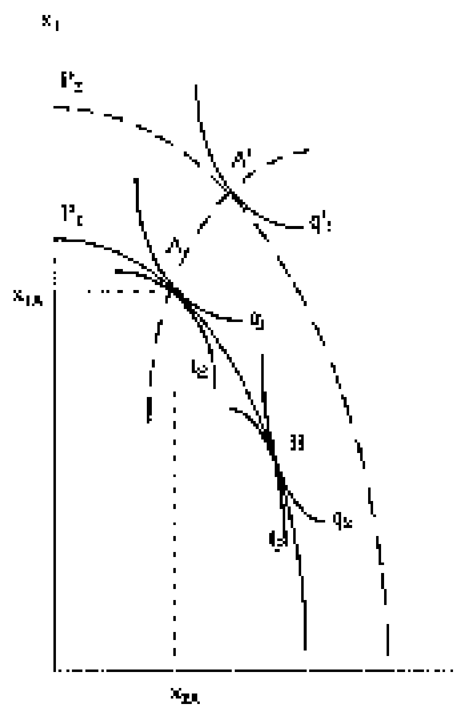
Buyers *l* and *k* in this market select the house type with characteristics that are closest to optimal. The point *A* represents the tangency of q_l and t_E with the hedonic price surface P_1 for consumer *l* and producer *E*, and the point *B* represents the tangency of q_k and t_F with the hedonic price surface for consumer *k* and producer *F*. The total expenditure on characteristics, the price of quality, is the slope of the hedonic surface above an expansion path such as *AA′* shown in Figure 1. Figure 1 also shows that housing types with different characteristics, though available at the same price, are chosen by different consumers. As shown, buyer *l* purchases house type *A* with characteristic level x_{1A} and x_{2A}. Rosen (1974) shows in markets with many buyers and sellers, the hedonic contours will trace out an envelope of tangencies between the bid and offer prices of the buyers and sellers. The realism of the nonlinear hedonic constraint requires that housing characteristics must be bought and sold in tie-in sales. We would expect tie-in sales for housing because housing characteristics cannot be unbundled from the geographic location of the house.

The discussion above and Figure 1 suggest that functional forms used to estimate hedonic prices should allow for the possibility of nonlinearity in the relationship between the price of the house and the prices and quantities of the underlying attributes. They also suggest that the divergence of tastes and technologies is an essential part of the theory of hedonic price functions and that "representative consumer" models may not describe market outcomes well.

The derivation of hedonic price functions outlined above views the price of houses as determined in a flow market—where housing supply comes from producers of housing

FIGURE 1

HEDONIC FRONTIERS



and price equilibrates the demand for new houses to the supply of new housing. An alternative view focuses on the stock of existing housing. In this case prices, again defined for attributes, guide both bids and offers for locational choices with respect to packages of housing characteristics (Alonso 1964, Muth 1969). The hedonic price function is determined by market clearing conditions in which the tie-in sales of attributes at each location equal the amount demanded by buyers. In equilibrium buyers and sellers are perfectly matched, and again the hedonic price surface is likely to be nonlinear.

The primary implication of the theoretical literature is that hedonic price functions are likely to be nonlinear because locational uniqueness leads to tie-in sales. Thus, observed housing prices reflect both the implicit prices of characteristics in housing packages and the quantities of characteristics embedded in the housing units sold. The theoretical structure of the market-clearing mechanisms for housing does not suggest that it can be assumed either that at a given market period the relative implicit prices for attributes are the same or that across market periods the implicit characteristics prices for the same packages of housing services remain constant. This inherent difficulty in interpreting observed housing price levels presents a particular problem for solving the index number problem for

frontier would be concave to the origin following classical utility theory, and in the latter case the frontier would be convex to the origin because it is a production transformation curve. Neither of these two cases is particularly helpful in the housing market since neither condition would be expected to be true.

housing—how to measure average price level changes across time periods.

II. EMPIRICAL HEDONIC PRICE INDEXES FOR HOUSING

A comprehensive review of the economic theory of index numbers and their use in housing markets is beyond the scope of this paper.[2] In brief, the index number problem for housing has much in common with the problem of index number construction for other goods and services. In a given base period, a "representative" consumer takes base period prices as given and buys a utility maximizing combination of goods and services, including housing services. In later periods, the consumer faces new sets of prices and selects alternative bundles of goods and services. The index number problem is to determine how much the cost of living has changed between periods if the consumer retains the original standard of living. The theory of hedonic price indexes for housing follows this literature, with the only modification being that economic agents select across composite characteristics.

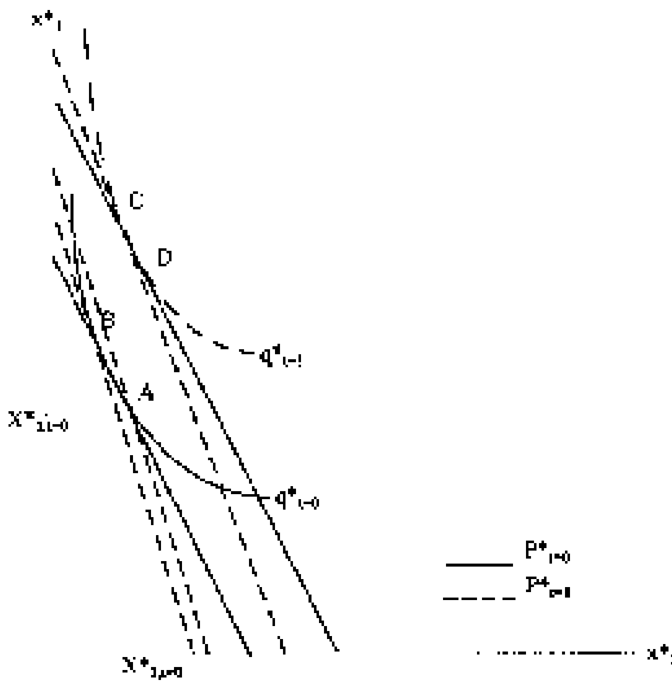
Figure 2 shows the price index problem for the more standard homogeneous goods case. Figure 2 shows optimal consumption in the two-commodity case, x_1^* and x_2^* over two periods, $t = 0$ and $t = 1$. As shown, there are several ways to measure relative price level changes. The first way, which is called Laspeyres price indexes,[3] holds the base period commodity bundle ($x^*_{1,\,t=0}$, $x^*_{2,\,t=0}$) fixed at point *A* and measures how much the base period bundle would cost at the subsequent period prices, $P^*_{t=1}$. The problem with the measure is that it does not account for the fact that at the new prices, $P^*_{t=1}$ the consumer would be expected to substitute to a new combination of goods, point *B*, while holding the level of well-being, or standard-of-living, $q^*_{t=0}$, constant. Because the Laspeyres price index is weighted on the initial bundle, point *A*, it does not account for the substitution effect and thus has an upward bias as a measure of the cost to the consumer of keeping the initial standard of living once prices have changed.

The alternative measure, the Paasche index, is similar except that it uses the subsequent period consumption bundle as its reference point, point *C*, for the subsequent period standard of living, $q^*_{t=1}$, and measures how much the subsequent period's consumption bundle would cost at the previous period's prices, $P^*_{t=0}$.[4] Here again, because the Paasche index weights on the $t = 1$ period's optimal consumption bundle, point *C*, it does not account for the substitution effect, point *D*, and thus has a downward bias as a measure of the cost to the consumer of keeping the $t = 1$ period standard of living, $q^*_{t=1}$. This bias arises because the bundle represented by point *C* was not the one actually chosen by the consumer in the base period, so computing its costs at the new prices overstates the cost of living in that period.

If one knew the consumer's preferences, either $q^*_{t=1}$ or $q^*_{t=0}$, one could measure the substitutions that would be made in order to maintain a constant level of well-being subsequent to a shift in relative prices for the two commodities. In fact, it would be possible to measure exactly

FIGURE 2

LASPEYRES AND PAASCHE PRICE INDEXES



2. For excellent discussions about the theory of index numbers and cost-of-living indexes, see Motley (1992), Pollak (1991), and Diewert (1983).

3. The general *n*-commodity Laspeyres price index measures the increase in prices from base period 0 to period *t* holding the initial consumption level constant:

$$Index_{Laspeyres} = \frac{\sum_{n=1}^{N} p^*_{nt}\, x^*_{n0}}{\sum_{n=1}^{N} p^*_{n0}\, x^*_{n0}}\ .$$

4. The general *n*-commodity Paasche price index measures the increase in prices from base period 0 to period *t* holding the t^{th} period consumption level constant:

$$Index_{Paasche} = \frac{\sum_{n=1}^{N} p^*_{nt}\, x^*_{nt}}{\sum_{n=1}^{N} p^*_{n0}\, x^*_{nt}}\ .$$

the difference in the minimum costs of obtaining any fixed level of satisfaction at any given set of prices. Such an exact cost-of-living index would be a measure of the true cost of maintaining a fixed level of satisfaction. The Laspeyres and Paasche indexes thus would be only approximations to the hypothetically exact cost-of-living indexes because they hold the observable consumption bundles fixed rather than the unobservable constant levels of satisfaction.

The economic theory of price indexes for heterogeneous goods such as housing follows the same logic as that of homogeneous goods, represented in Figure 2 (Triplett 1987, 1989). Instead of considering the consumer's optimal consumption combinations of commodities, we would consider their optimal consumption combinations of characteristics of housing. This implies that the axes x_1^* and x_2^* of Figure 2 should be redefined as composite characteristics of the housing asset, and the budget constraints, $P^*_{t=0}$ and $P^*_{t=1}$ should be redrawn as nonlinear functions. Construction of approximate hedonic cost-of-living indexes (or more appropriately sub-indexes) would then proceed analogously to the homogeneous goods framework. However, now both the preferences, $q^*_{t=0}$ and $q^*_{t=1}$, and the true nonlinear hedonic surfaces are unobservable.[5]

Empirical estimates of the hedonic price function can be obtained for alternative price regimes, and these can be evaluated using either fixed characteristics weights from the beginning period, a fixed-weight Laspeyres-type index, or using fixed characteristics weights from the end of the period, a fixed-weight Paasche-type index. Price indexes obtained in this manner can be interpreted as approximations to the exact cost-of-living index for housing. They are approximations in the sense that they contain only information about the hedonic at a fixed set of characteristics between two time periods, whereas the true indexes also require information about preferences or levels of satisfaction. The Laspeyres-type and Paasche-type cost-of-living indexes for housing will therefore suffer from the same substitution bias found in their counterparts for homogeneous goods. The Laspeyres-type housing index would be expected to be biased upward and the Paasche-type index would be expected to be biased downward (Diewert 1983).

Triplett (1987) speculates, though does not prove, that empirical hedonic-index approximations may provide bounds on the true characteristics price index in the same way that the Laspeyres and Paasche indexes do in the homogeneous goods case. Diewert (1978) argues that if the empirical Laspeyres and Paasche indexes lie "close" to each other then the Fisher Ideal index[6] should be "close" to a reference exact price index that lies between the exact Paasche and Laspeyres price indexes.

Although there appear to be a number of similarities between the hedonic approximations for the empirical Laspeyres and Paasche cost-of-living indexes and those obtained for homogeneous goods, Triplett (1987) argues that there are also important differences. First, the form of the hedonic surface (the implicit prices of the characteristics) must be estimated empirically and, other than nonlinearity, there are no theoretical guidelines about appropriate functional forms. Second, the usual statistical procedures produce estimates for a shift in the whole hedonic surface rather than an estimate for shifts in a single selected budget hyperplane (the shift in the prices holding characteristic levels constant) as required in the fixed-weight empirical indexes.

Another problem is the goodness of the approximation. The empirical index numbers, such as the Laspeyres, Paasche, or Fisher's Ideal, use only price and quantity information, not the unobservable preferences. Thus, they are approximations to the theoretically correct, or exact, index numbers because they only approximately hold utility constant over the index comparison periods. With approximations, an error of indeterminable size is introduced into the index every time the fixed utility assumption is violated by changes in relative characteristics prices. As discussed, recent empirical and theoretical work indicates that good approximations to exact indexes can be computed from fixed-weight formulae. Thus, the criterion for the "goodness" of these index approximations in empirical applications is the extent to which the computed index takes account of, and controls for, variation in housing characteristics or quality. Quality variation is measured as the characteristics sets that are embodied in the housing stock from period to period. The fixed-weight approximations must fix these characteristics sets at either the beginning or end of the analysis period.

Several conclusions from cost-of-living index theory have practical implications for the empirical task of constructing housing price indexes. Pollak (1991, p. 168) suggests that it is useful to view the theoretical implications by distinguishing between the "estimation stage" concerning the appropriate specification of the hedonic price function (equation (1) above) and the "composition stage" in

5. The nonlinearity of the hedonic boundary constraint invalidates the usual strategy used for constructing cost-of-living indexes for homogeneous goods, in which it is assumed that the budget constraint (defined in consumption goods space) is a bounding hyperplane whose linearity assures that there is a duality between the utility function and the consumption cost function.

6. The Fisher Ideal price index is defined as the geometric average of the Laspeyres and Paasche price indexes

$$Index_{Fisher} = \sqrt{\left(Index_{Paasche}\right) \cdot \left(Index_{Laspeyres}\right)}\ .$$

which the estimated hedonics are used to obtain price indexes. For the estimation stage, it was shown above that the hedonic function is, in Rosen's terminology, an estimate of the minimum price of any package of characteristics (Rosen, 1974, p. 37) and thus, it is the empirical counterpart to the characteristics cost function. In a market such as housing, with a continuous variety spectrum, the functional form for the hedonic is an empirical question. In general, however, the characteristics price (the partial derivatives of the characteristics cost function) are themselves functions whose value depends on the particular point in the characteristics space where they are evaluated. This suggests that empirical specifications should allow for maximum flexibility of functional form. Theory also has little to say about the elements of the characteristics set used to estimate the hedonic. Theoretically, the chosen set should include all characteristics that can reasonably be assumed to enter household preferences.[7] Finally, it would be desirable to use estimation strategies that provide local approximations to the hedonic price function at fixed characteristics levels in each time period. In this way, it would be possible to control for a fixed consumption bundle and obtain better estimates for either the Laspeyres or the Paasche index approximations.

The second implication of the theory concerns the "composition stage" of the price index. Once an empirical estimate of the hedonic is obtained, what is the appropriate composition of the price index? It was argued that the theoretically exact index could not be uncovered due to the nonlinearity of the hedonic and lack of information about preferences. Thus, suitable approximations are measures of the effects of relative price changes when the beginning point, or end point, of the characteristics bundle is fixed. This strategy ignores the substitution effects from price changes. The practical empirical task is to obtain estimates of the hedonic price surface such that unbiased estimates of the prices of fixed sets of characteristics can be computed.

III. ESTIMATING HEDONIC-BASED PRICE FUNCTIONS

The primary theoretical objectives for the estimation of hedonic housing functions are that the estimation strategy should allow for the nonlinearity of the hedonic contours and that it should provide an accurate accounting of, and control for, variations in characteristics, or quality, over time. The primary criticism that has been raised against the hedonic methodology concerns the appropriate way to meet these theoretical objectives in the usual regression framework. The first complaint is that the "correct" set of characteristics must be selected to achieve an unbiased estimate of the hedonic function. The second complaint is thata priori assumptions concerning the "correct" functional form must be imposed to estimate the hedonic function in a regression framework. A final complaint is that hedonic price function estimates are likely to suffer from sample selection bias because they are obtained from samples of transactions that may not be random samples of the population of house prices.

Alternative Specifications to Control for Characteristics

An important alternative recommended strategy is the repeat sales methodology, which was first introduced by Bailey, Muth, and Nourse (1963) and further developed by Case and Shiller (1987, 1989). This method focuses on price changes rather than price levels, and it restricts estimation to a subsample of houses that have not changed their characteristics set and have sold at least twice. The primary advantage of this strategy is that it avoids the specification of the characteristics set for houses and the functional relationship of characteristics to price. The argument is that first differencing the log of house prices and using only houses that have been sold at least twice and have not changed their characteristics produces a perfect control for the entire set of relevant characteristics.

The primary advantage of the repeat sales methodology also imposes important theoretical restrictions on the admissible class of characteristics cost functions that can be considered. It can also be shown, (Meese and Wallace 1996, Wang and Zorn 1995) that the estimated coefficients in the repeat sales framework are complicated frequency weightings of the simple means of logarithm of the ratio of final transaction price to the initial transaction price over relevant time periods.[8] These weights do not have a "fixed-weight" interpretation in the sense discussed above because

7. It is this point that Shiller (1993) identifies as the greatest weakness of the housing price indexes composed from hedonic price function estimates. He argues that these decisions are necessarily arbitrary because they involve "...not only the decision of which quality variable to include, but there are also decisions to make about allowing nonlinear effects of each and interaction effects..."(p. 129). He also asserts that the lack of available characteristics data leads to problems with sample size and misspecification due to omitted characteristics.

8. For example, in a three-period sample with possible repeat sales between periods 1 and 2, 1 and 3, and 2 and 3, the least squares estimators for the logarithm of the index number for periods 2 and 3, respectively, are:

$$\hat{\phi}_2 = \frac{n_{12}(n_{13}+n_{23})\bar{r}_{12} + n_{13}n_{23}(\bar{r}_{13} - \bar{r}_{23})}{n_{12}(n_{13}+n_{23}) + n_{13}n_{23}}$$

the computed means reflect different subsamples of unobservable characteristics bundles. Thus the estimated price relatives do not provide an estimate of the characteristics cost frontier at a fixed package of characteristics as required in the usual formulation of approximations to exact cost-of-living price indexes. It has this interpretation only if it is assumed that the true hedonic contours shrink toward the origin in a homogeneous fashion.

The repeat sales strategy also assumes that the characteristics levels for houses do not change and those that do can be "correctly" identified. This assumption leaves the measure vulnerable to the same misspecification concerns that the hedonic methodology must contend with. Finally, the repeat sales method requires careful testing of sample selection assumptions because the sample is by definition more restrictive than those used in the hedonic methodology.[9] These trade-offs suggest that further refinements may be required for both methods. These refinements include development of hybrid methods that combine features of both methods (Quigley 1995).

Flexible Nonparametric Estimation Strategies

Flexible nonparametric estimation strategies directly address the problem of imposing a priori specifications on the hedonic functional forms or using grid search methods over a limited class of functional forms. They also allow local approximations to the hedonic surface at fixed points, which is more in keeping with the requirements of price index formation.

Following Meese and Wallace (1991), suppose the natural log of house price in period t, $P(n,t)$, varies with the natural log of its characteristics, $x(n,t)$, according to a hedonic function:

$$(4) \qquad P(n,t) = m(t) + \beta_t' G[x(n,t)] + u(n,t),$$

where $m(t)$ accounts for the changing residual mean in house prices, β_t denotes a $(j \times 1)$ vector of parameters, $x(n,t)$ is a set of j housing characteristics observed for the nth transaction at time t, G is a function of the characteristics, and $u(n,t)$ is an additive error term. The nonstationary mean in housing prices is attributed to the drift, $m(t)$, which is modeled as:

$$(5) \qquad m(t) = \alpha(t) dum(t)$$

where $dum(t)$ is a dummy variable equal to one for each quarterly observation period t and zero otherwise, $\alpha(t)$ is the regression parameter measuring period t residual mean price change between periods once the mean changes in characteristics costs have been accounted for, and $e(t)$ is the time-series error component that is assumed to be white noise. Combining (4) and (5) yields a fully general hedonic function:

$$(6) \qquad P(n,t) = \beta_t' G[x(n,t)] + \alpha(t) dum(t) + (e(t) + u(n,t)).$$

From a theoretical perspective, the preferred method to estimate equation (6) is a strategy that imposes the fewest a priori restrictions on the functional form of $G[\cdot]$.[10] Nonparametric methods allow for the greatest possible flexibility in estimating functional forms and allow empirical estimation of data contours over a wide range of smooth functions. A particularly suitable nonparametric method is regression by *loess* which was first introduced by Cleveland and Devlin (1988) and Cleveland, Devlin, and Grosse (1988). *Loess* also allows for local approximations to the $G[\cdot]$ function at fixed points in the data surface.

Loess is a technique for estimating a regression surface in a moving average manner and can approximate a wide range of smooth functions. Meese and Wallace (1991, 1996) use a version of the regression model in equation (6):

$$(7) \qquad P(n,t) - P(mean,t) = \beta_t' G[X(n,t)] + v(n,t), \quad n = 1,\ldots,N(t),\ t = 1,\ldots,T$$

where $P(mean,t)$ is the quarterly mean of the logarithm of housing prices and $v(n,t)$ is the composite error term. Because nonparametric local fitting strategies require stationary dependent and independent variables, I remove the trend in $P(n,t)$ by subtracting the quarterly mean of the dependent variable each quarter and then standardize the variable by dividing by the quarterly sample standard deviation. I also standardize all the characteristics variables by subtracting the global mean and dividing by the sample standard deviation.

$$\hat{\phi}_3 = \frac{n_{13}(n_{12}+n_{23})\bar{r}_{13} + n_{12}n_{23}(\bar{r}_{12}+\bar{r}_{23})}{n_{12}(n_{13}+n_{23}) + n_{13}n_{23}}$$

where $\bar{r}_{12}$, $\bar{r}_{13}$, and $\bar{r}_{23}$ are the means of the logarithm of the ratio of final transaction prices to the initial transaction prices in the subscripted time interval, and n_{12}, n_{13}, and n_{23} are the sample frequencies for repeat sales in the subscripted time interval.

9. Meese and Wallace (1996) test the repeat sales assumption that the characteristics prices are time-invariant using a second order Taylor series approximation to the hedonic function and a transaction data set from Alameda County. They reject the assumption for all municipalities, suggesting that this is not an innocuous maintained hypothesis.

10. Meese and Wallace (1991) test for parametric flexible functional forms such as the translog and the log-log function as do Halvorsen and Pollakowski (1981). Their findings suggest no consistent preference for one specification across municipalities.

I employ two centering strategies. The first strategy I call the nonfixed-centering *loess* estimator. It uses the vector of mean characteristics for quarter t, $X(m,t)$, to center the local fitting of $G(\cdot)$. *Loess* uses a fraction n^*, $0 < n^* < 1$, of the total number of observations closest to $X(m,t)$, where proximity is measured using the Euclidean distance between all points in the sample and $X(m,t)$. The distance metric is defined by:

$$\text{(8)} \qquad D[X(m,t), X(n,t) = [\Sigma X(m,t) - X(n,t)]^2]^{1/2},$$

where the summation runs over the j-elements of the set of housing characteristics.

The hedonic surface is approximated locally at $X(m,t)$ by a weighted least squares regression for the n^* observations nearest $X(m,t)$. The weights are defined by Cleveland and Devlin (1988) as:

$$\text{(9)} \qquad W = V[D(X(m,t)\ X(n,t)) / D(X(m,t), X(n,*))],$$

where $D(X(m,t), X(n,*))$ is the distance from the mean X in a given quarter to its n^* nearest neighbors. Following Cleveland and Devlin (1988) and Meese and Wallace (1991) I use the "tricube" functional form for $V[\cdot]$.[11] This strategy provides estimates for the curvature of the hedonic price function at the mean characteristics over the 101 quarters in the data set.[12] I set n^* at 0.33 in an effort to balance the trade-off between bias and sampling error.

There are two problems with this strategy. The first is that the quarterly means are used to detrend the price data and these means reflect both price changes and changes in the set of characteristics traded. The second problem is that *loess* is estimated by centering at the quarterly means, whereas the desired price estimates should be centered at a fixed characteristics set from the beginning or end of the period.

The second strategy addresses these problems. The fixed-centering *loess* estimator centers at two fixed characteristics sets: the first set is fixed at the mean, $X(m_L,1)$, of the characteristics for 1970:Q1, a Laspeyres-type estimator, and the second set is fixed at the mean, $X(m_P,101)$, of the characteristics for 1995:Q1, a Paasche-type estimator.[13] Thus, these estimators replace the term $X(m,t)$ with the appropriate fixed characteristics set. The estimation is then carried out for the n^*_Q nearest neighbors for each quarter using the same weighting strategy as in the non-fixed centering strategy.

There remains one problem with the fixed-centering *loess*. For smaller municipalities, the *loess* weightingscheme leads to an insufficient number of observations in the neighborhood of the initial characteristics set in some quarters. It is thus necessary to smooth across quarters, although in the applications here one never has to smooth over more than two quarters. The primary advantage of the fixed estimator is that it is consistent with the requirements of empirical Laspeyres and Paasche-type price indexes.

The transaction data used in this analysis included four characteristics: number of bathrooms, number of bedrooms, square footage of the living area, and the age of the dwelling. I constructed a variable bedrooms/living area to account for possible nonlinearities from adding more bedrooms onto a home of a given square footage. Because homes with a high ratio of bedrooms to living area are likely to be rental property, often for student habitation, I expected that higher ratios would reduce house prices. The other characteristics, except for the age of the dwelling, were expected to have positive effects on housing prices. I did not have strong priors on the effect of age on house price. The age variable may well proxy for other unmeasured features of the dwelling such as architectural design. For example, many older California craftsman homes sell at a premium due to their distinctive design characteristics; on the other hand, age could account for the effects of deterioration or a lack of modern room organization.

The results for the nonfixed and fixed-centering *loess* estimates are reported in Tables 1 and 2.[14] The price elasticities are obtained by taking the derivative of the estimated housing price function with respect to each characteristic and evaluating the derivative at the appropriate characteristics set. The Tables report two types of elasticities for 1970:Q1 and for 1995:Q1. Reading down the columns for each municipality, the Laspeyres-type elasticities are evaluated at the mean characteristics set for the first quarter of 1970 for each municipality. The Paasche-type elasticities

11. The tricube $V[\cdot] = (1 - s^3)^3$, if $s < 1$; it is equal to 0 otherwise. The advantage of the tricube is that it allows smooth contact with 0 and 1 endpoints.

12. The 1970 through first quarter 1988 Alameda County data were obtained from the California Market Data Cooperative and account for about 98% of all arm's-length transactions over the period. The 1988 through 1995 data were obtained from Property Sciences, Inc. and TRW.

13. This assumes that the observed quarter 1 sample is a random sample of the characteristics set for houses traded in 1970 and similarly for the 1995:Q1 sample.

14. The skewness and kurtosis measures for the residual distributions from these estimates have close to symmetric distributions, although they have fatter tails than would be expected under normality. The White test for heteroskedasticity in the residuals indicates that there remains contemporaneous heteroskedasticity for several of the municipalities. These diagnostics suggest that the more efficient estimates of the characteristics prices should be considered. A dynamic model might include allowance for serial correlation and/or ARCH in the time-series component of the composite error, or explicit consideration of the speed of adjustment of prices to changes in market fundamentals or levels of housing characteristics.

TABLE 1

ESTIMATED PRICE ELASTICITIES FOR THE NONFIXED-CENTERING *LOESS* ALAMEDA COUNTY MUNICIPALITIES: 1970:Q1–1995: Q1

ALAMEDA COUNTY	BATHROOMS		BEDROOMS/ TOTAL LIVING AREA		TOTAL LIVING AREA		AGE OF HOUSE	
	Paasche	Laspeyres	Paasche	Laspeyres	Paasche	Laspeyres	Paasche	Laspeyres
ALAMEDA[3]								
1970	9,563	1,210	3,995	520	126	12	740	112
1995	9,699	1,228	3,995	501	126	12	744	113
ALBANY								
1970	20,477	1,294	36,354	5,576	179	18	2,350	–1,026
1995	36,860	2,330	–7,383	1,132	155	16	307	133
BERKELEY								
1970	7,560	763	–8,836	–1,211	199	19	827	100
1995	11,190	1,129	–14,073	–1,929	203	20	727	89
CASTRO VALLEY[1,2,3]								
1970	7,947	1,404	4,681	9,521	40	8	1,290	311
1995	17,118	3,024	4,057	825	68	13	1,402	339
DUBLIN[1,2,3]								
1970	18,470	2,814	–2,463	–361	88	12	2,414	831
1995	25,574	3,897	–36,041	–5,292	167	23	4,161	1,433
HAYWARD[1]								
1970	7,239	1,233	–2,614	–416	75	11	980	238
1995	13,444	2,290	–5,809	–926	93	14	1,004	243
FREMONT[1,2,3]								
1970	12,356	1,310	–5,734	–749	70	10	1,171	259
1995	22,582	2,395	1,720	225	77	11	1,273	282
LIVERMORE[1,2,3]								
1970	10,389	1,298	3,538	649	42	8	1,717	306
1995	13,961	1,745	1,927	353	53	18	1,787	319
NEWARK[1]								
1970	12,788	1,099	3,444	470	13	1	2,260	441
1995	33,154	2,849	–898	–122	61	6	3,002	585
OAKLAND								
1970	15,933	2,418	–574	–98	119	20	1,248	204
1995	32,432	4,924	3,148	541	121	21	–365	–59
PIEDMONT[1,2,3]								
1970	58,737	3,565	–2,588	–135	105	10	2,896	101
1995	136,857	8,306	–48,785	–2,560	199	19	–5,494	–192
PLEASANTON[1,3]								
1970	36,939	4,439	80	974	121	16	1,487	1,449
1995	49,514	5,947	15,382	1,852	151	20	694	676
SAN LEANDRO								
1970	8,001	1,551	988	304	40	7	765	161
1995	10,784	2,091	751	155	92	16	–655	–138
SAN LORENZO[2]								
1970	7,344	1,327	–6,670	–661	5	1	859	127
1995	9,487	1,714	6,107	1,025	38	8	–3,390	–225
UNION CITY[1]								
1970	9,695	1,326	–6,670	–661	5	1	859	127
1995	16,967	2,321	–3,174	–315	9	2	1,364	202

1. Statistically significant at 5% level, positive trend in bathrooms.
2. Statistically significant at 5% level, positive trend in bedrooms.
3. Statistically significant at 5% level, positive trend in living area.

TABLE 2

ESTIMATED PRICE ELASTICITIES FOR THE FIXED-CENTERING *LOESS* ALAMEDA COUNTY MUNICIPALITIES: 1970:Q1–1995: Q1

ALAMEDA COUNTY	BATHROOMS		BEDROOMS/ TOTAL LIVING AREA		TOTAL LIVING AREA		AGEOF HOUSE	
	Paasche	Laspeyres	Paasche	Laspeyres	Paasche	Laspeyres	Paasche	Laspeyres
ALAMEDA								
1970	11,384	692	135	–1,694	109	11	1,053	–122
1995	16,394	1903	–9,480	–891	107	10	1,340	203
ALBANY[3]								
1970	9,829	156	108	–172	122	16	–456	154
1995	44,232	3,262	749	344	127	13	–124	–97
BERKELEY								
1970	13,105	712	3,600	–10,991	162	16	1,204	142
1995	25,203	3,058	–295	224	126	12	2,668	240
CASTRO VALLEY[1,2,3]								
1970	10,087	216	–4,486	–515	48	9	2,258	389
1995	11,310	378	3,765	1,309	77	14	864	159
DUBLIN[1,2,3]								
1970	7,104	584	–2,225	–2,132	71	7	–574	–205
1995	28,416	2,923	–10,112	–1,942	50	16	1,242	617
HAYWARD[1]								
1970	4,653	150	8,714	1,466	68	9	1,578	258
1995	11,634	748	726	1,041	71	12	514	23
FREMONT[1,2,3]								
1970	22,156	2,124	–1,762	337	87	12	–886	–252
1995	18,321	2,395	–5,447	824	90	13	–920	–189
LIVERMORE[1,2,3]								
1970	12,337	1,055	–3,660	1,388	57	7	964	52
1995	15,260	1,371	–5,856	–895	72	15	164	43
NEWARK[1]								
1970	12,314	1,628	4,192	–1,104	99	5	873	441
1995	14,682	2,320	898	–286	98	7	1,457	218
OAKLAND								
1970	14,311	2,558	1,481	–95	66	11	918	144
1995	22,898	3,234	–1,297	–445	66	12	739	115
PIEDMONT[1,2,3]								
1970	43,378	1,329	–44,802	–3,866	114	8	1,536	156
1995	100,103	2,172	57,248	1,593	218	19	1,769	–132
PLEASANTON[1,3]								
1970	13,143	849	–1,596	–1,462	112	13	–553	–316
1995	19,649	4,625	–2,264	–564	108	13	623	–2,526
SAN LEANDRO								
1970	10,784	1,146	2,819	4,272	76	12	940	163
1995	18,697	1,281	–1,691	–7	71	13	1,257	221
SAN LORENZO[2]								
1970	9,793	1,382	3,327	1,876	12	7	12	68
1995	12,548	1,880	11,645	1,916	48	11	–1,629	–121
UNION CITY[1]								
1970	5,508	884	–1,118	–239	66	14	1,784	121
1995	8,080	1,824	–11,771	–955	128	32	–2,410	420

1. Statistically significant at 5% level, positive trend in bathrooms.
2. Statistically significant at 5% level, positive trend in bedrooms.
3. Statistically significant at 5% level, positive trend in living area.

are evaluated at the mean characteristics set for the first quarter of 1995 for each municipality.

For the non-fixed centering *loess* reported in Table 1, the parameter estimates underlying the Paasche-type and Laspeyres-type elasticities are the same for each year (e.g., 1970:Q1 has one set of estimates and 1995:Q1 another).[15] Thus, the differences in the magnitudes of the elasticities come from the growth in the attribute sets from 1970:Q1 and 1995:Q1. The parameter estimates for the Paasche-type and Laspeyres-type elasticities are estimated separately for the fixed-centering *loess*. Thus, these elasticities reflect both changes in prices and growth in the characteristics set over the analysis period. The footnotes indicate whether there was a statistically significant trend in the mean levels of characteristics for bathrooms, bedrooms, and living area over the quarters. As shown, ten of the fifteen municipalities experienced statistically significant positive trend in the mean levels of these characteristics over the 101 quarters. Additionally, as expected in some municipalities, increasing the ratio of bedrooms to total living area reduces the value of the house. There is, however, quite a lot of variability in this result across the municipalities. The effect of age also varied across the municipalities; however, for most municipalities, increasing the age of the dwelling led to increases in housing prices.

The differences between the Paasche-type and Laspeyres-type elasticities by characteristics by municipality reflect changes in the magnitudes of the mean level of the characteristics set between 1970:Q1 and 1995:Q1. The nonfixed-centering *loess* estimation reported in Table 1, however, does not account for differences in the coefficient estimates at different mean levels of characteristics on the hedonic surface within a quarter. In Table 2, however, the fixed-centering *loess* estimates provide a local approximation to the hedonic at either the fixed 1970:Q1 characteristics level or the fixed 1995:Q1 level. Thus, the Table 2 elasticities control for the growth in the mean value of characteristics over the quarters, the changes in price levels of mean characteristics across quarters, and the differences in price levels within a quarter for different mean characteristics levels. The Table 1 elasticities control for only the growth in the mean value of characteristics over the quarters and the changes in price levels of mean characteristics across quarters. They do not control for differences in mean price levels within each quarter for different characteristics bundles.

For example, Castro Valley has experienced considerable growth in the mean levels of characteristics in houses sold from 1970:Q1 to 1995:Q1; the nonfixed-centering *loess* Laspeyres-type and Paasche-type price elasticities for bathrooms indicate about a 115% increase in the elasticities over the analysis period. The fixed-centering *loess* elasticities reported in Table 2, in contrast, indicate that the Paasche-type elasticity, holding the characteristics mean fixed at 1995:Q1 levels, experienced only a 12% increase and the Laspeyres experienced only a 75% increase from the 1970:Q1 mean level of characteristics. Similar differences appear in the elasticity of square footage. Fremont and Piedmont also experienced growth in mean characteristics levels over the period. Here again, the price elasticity for bathrooms increased by 82% for Fremont using the Table 1, nonfixed-centering *loess* results, whereas the price elasticity of bathrooms fell by 17% using the Paasche-type fixed *loess* estimates. The Piedmont elasticity of bathrooms increased by 132% using the Table 1 estimates, however, the elasticity growth found for the Table 2 fixed estimates was between 131% and 63%. The results for the square footage elasticities were similar. The elasticity results for the ratio of bedrooms to total rooms is similar in many municipalities, although it is difficult to interpret the negative changes in Livermore. A reasonable conclusion from comparing Tables 1 and 2 is that the differences in the results are most pronounced for the municipalities that experienced the most growth in the mean levels of the characteristics, such as Castro Valley, Dublin, Fremont, Livermore, and Piedmont.

Oakland did not experience statistically significant growth in the mean level of characteristics over the period; however, there is also evidence of the effects of confounding characteristics level growth with price changes. The Oakland price elasticity for bathrooms increased about 104% using the Table 1 estimates, whereas it grew only 60% for the Paasche-type elasticity and 26% for the Laspeyres-type elasticity. Thus, even in a municipality in which the growth of the mean characteristics was not sustained there appears to be confounding of the growth in the mean levels of characteristics with changes in the relative price levels. The fixed *loess* results appear to control better for the confounding effects of growth in the level of the characteristics.

15. To reiterate, the difference between the two estimation strategies is that the non-fixed centering *loess* uses the mean level of characteristics in each quarter and then selects the nearest neighbors from all the data,whereas the fixed centering *loess* estimation obtains two estimates: one centered at the mean of the 1970:Q1 characteristics and the other centered at the mean of the 1995:Q1 characteristics, and the nearest neighbor is determined within a quarter. The price elasticities are then obtained using the coefficients for the mean initial and end-of-period characteristics set. The fixed estimation evaluates the elasticities for the beginning quarter and ending quarter coefficients using either the first quarter mean characteristics (a Laspeyres-type measure) or the last quarter mean characteristics (a Paasche-type measure).

I conclude that the hedonic price surfaces can consistently be estimated with both *loess* strategies, although the fixed strategy is somewhat more consistent with the theoretical structure of empirical Laspeyres and Paasche price indexes. The most important difference between the two strategies is found for the characteristics price for housing attributes that have changed the most over the 25-year period. Finally, the “hedonic” or characteristics effects account for a substantial part of the change in house prices.

IV. CONSTRUCTING HOUSING PRICE INDEXES

As previously discussed, consistent estimates of the hedonic surface can be used to construct estimates of the Laspeyres-type and Paasche-type price indexes. The theoretically desirable Fisher Ideal price index can be computed from the geometric average of these two bounds. As Diewert (1978) has shown, if the Laspeyres-type and the Paasche-type price indexes are very close to one another, the Fisher Ideal can be considered as a close approximation to an exact price index defined in characteristics. The usual sense in which price indexes are considered to be close approximations relates to the degree to which they control for fixed levels of characteristics in the construction of the price index. An advantage of the fixed-centering *loess* estimation is that it allows for local approximations to the hedonic price at fixed mean levels of characteristics. Thus, the fixed-centering *loess* seems to be the preferable estimation strategy given the empirical results summarized in Tables 1 and 2 and the theoretical requirements for close approximation strategies for index number construction.

Figures 3–8 compare the fixed and nonfixed-centering *loess* Fisher Ideal price indexes with repeat sales indexes, quarterly means, and quarterly medians for three municipalities: Oakland, Fremont, and Piedmont. Oakland experienced relatively little growth in the mean level of housing characteristics over the analysis period and Fremont and Piedmont experienced considerable growth in mean housing characteristics. Figure 3 compares the fixed-centering *loess* Fisher Ideal price index with the quarterly means of house prices and a repeat sales price index for Oakland. The quarterly means exceed the fixed Fisher Ideal index and the repeat sales index for nearly all the quarters. Repeat sales accounted for only 19% of all sales over the sample period, and the repeat sales index appears to underestimate the price index consistently. The fixed Fisher Ideal appears to account for the confounding effects of the mean levels of characteristics from the changes in relative prices of the characteristics. Figure 4 is consistent with the results in Tables 1 and 2 in that the fixed Fisher Ideal shows a smaller

FIGURE 3

OAKLAND: FIXED *LOESS* FISHER IDEAL, MEAN, AND REPEAT SALES PRICE INDEXES

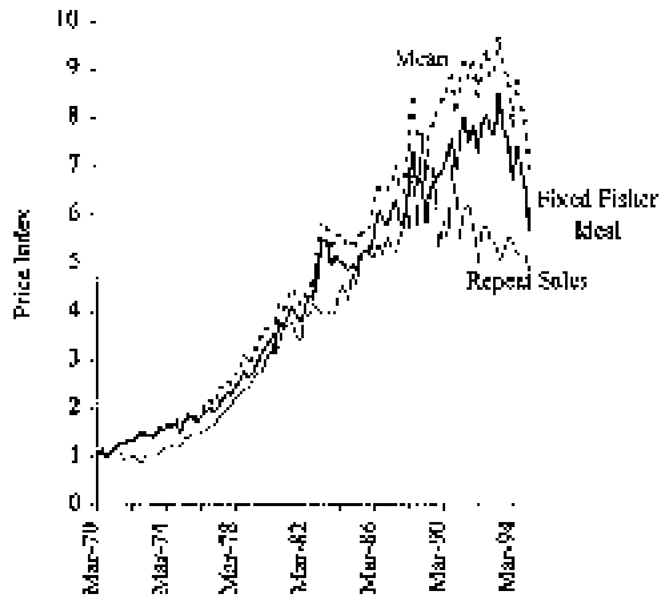


FIGURE 4

OAKLAND: FIXED AND NONFIXED CENTERING *LOESS* FISHER IDEAL PRICE INDEXES

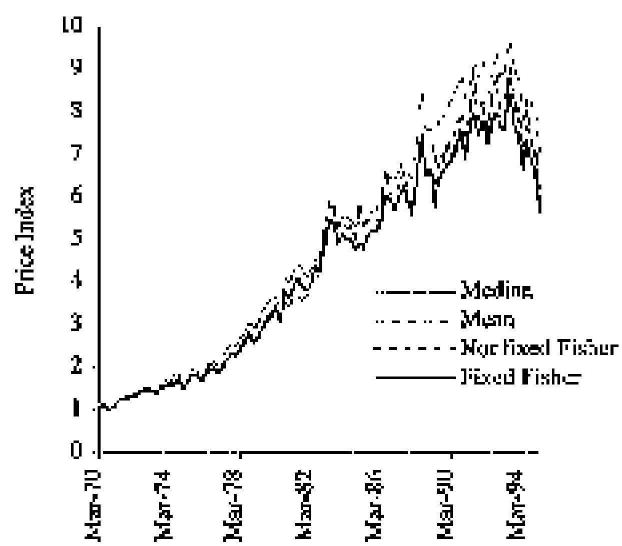


FIGURE 5

Fremont: Fixed *Loess* Fisher Ideal, Mean, and Repeat Sales Price Indexes

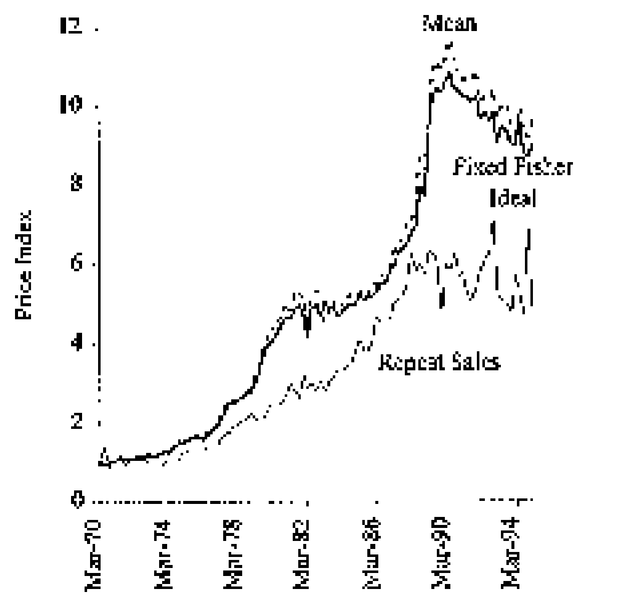


FIGURE 7

Piedmont: Fixed *Loess* Fisher Ideal, Mean, and Repeat Sales Price Indexes

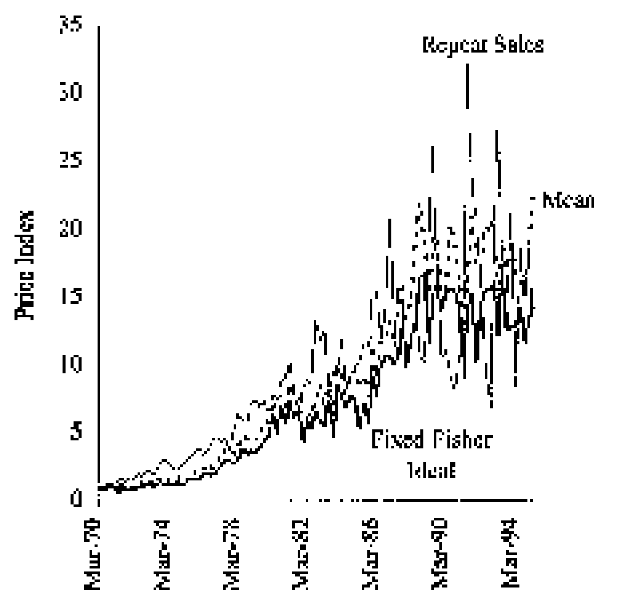


FIGURE 6

Fremont: Fixed and Nonfixed Centering *Loess* Fisher Ideal Price Indexes

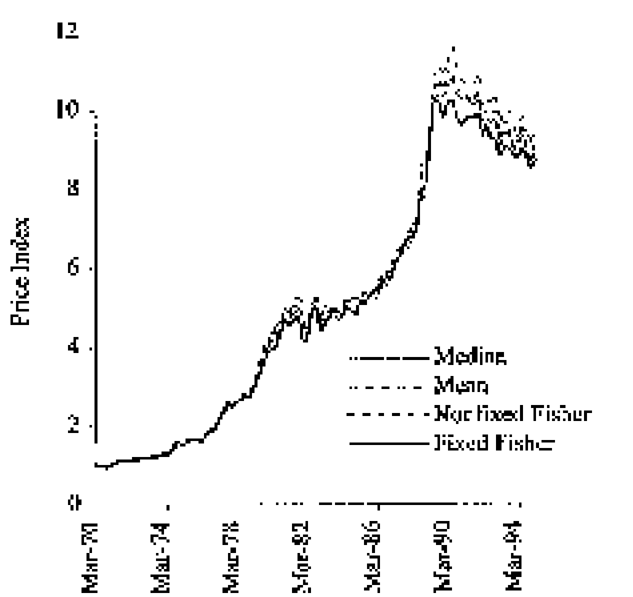
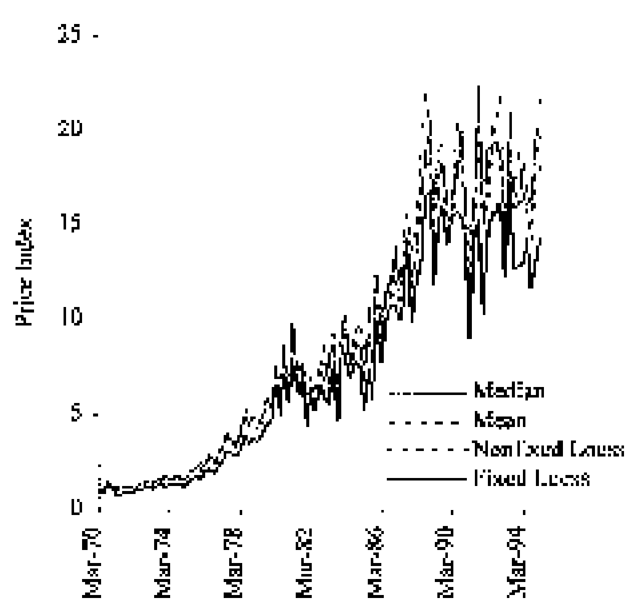


FIGURE 8

Piedmont: Fixed and Nonfixed Centering *Loess* Fisher Ideal Price Indexes



change in relative prices than either the quarterly means or the nonfixed-centering Fisher Ideal index. The differences between the two Fisher Ideal indexes are important only after the first quarter of 1989, when the fixed-centering Fisher Ideal falls below the nonfixed-centering index.

Figures 5 and 6 provide the same information for Fremont. The results are similar to the Oakland graphs, although the quarterly means more closely track the fixed-centering Fisher Ideal. The repeat sales index again substantially underestimates the relative price changes compared to the quarterly means and the fixed-centering *loess* Fisher Ideal. Repeat sales account for about 18% of the total sales over the period in Fremont. Figure 6 compares the fixed and nonfixed-centering *loess* Fisher Ideals with the quarterly mean and median indexes. The fixed Fisher Ideal is consistently below the nonfixed Fisher Ideal, as expected from the results of Tables 1 and 2.

Figures 7 and 8 provide the index construction results for the city of Piedmont. Piedmont is an exclusive residential community that is entirely surrounded by the city of Oakland, but all its public service systems, including schools, are separate from those of Oakland. Piedmont has experienced growth in the mean levels of characteristics of housing sold during the period as well as very substantial price appreciation of attributes. Figure 7 compares the repeat sales index with the quarterly mean index and the fixed-centering Fisher Ideal. Again, the quarterly mean index appears to overestimate the appreciation of house prices. The repeat sales index is wildly erratic, most probably due to the small sample size for repeat sales in Piedmont, only 630 homes. The fixed Fisher Ideal index appears to control for the confounding effects of the growth in the mean levels of characteristics and is considerably less erratic, due to the larger sample size. Figure 8 compares the fixed and nonfixed-centering Fisher Ideal indexes with the quarterly median and mean indexes. Again the fixed Fisher Ideal lies everywhere below the nonfixed index, which more closely tracks the mean and median indexes.

These graphical results appear to indicate that accounting for the growth in the mean levels of characteristics gives a rather different view of house price increases in Alameda County municipalities. The fixed-centering *loess* Fisher Ideal index is particularly appealing because it allows local approximations of the hedonic surface as prescribed by the theory of cost-of-living indexes and allows for the construction of Fisher Ideal price indexes that are the geometric average of the beginning and ending period characteristics levels. The elasticities derived from the estimation of the hedonic appear to suggest that fixing the point of approximation may be necessary to avoid confounding the growth in the levels of characteristics, which can be viewed as a measure of quality, from the changes in the relative prices of characteristics. The results from comparing the constructed Fisher Ideal indexes also indicate that the fixed approximation may be preferable.

V. CONCLUSIONS

This paper reviewed basic principles of price index construction for heterogeneous goods such as housing, where differing levels of characteristics (quality) lead to important differences in prices. The price/quality relationship is described by the housing price hedonic, which is likely to be nonlinear. Nonparametric econometric techniques are particularly suitable for the hedonic price function estimation problem because they allow for many classes of functional forms. I show how one nonparametric technique, *loess*, allows for the added feature of centering the estimation to fixed points, such as the beginning or ending period characteristics sets consistent with the requirements of Laspeyres-type and Paasche-type price indexes.

The *loess* estimates for the hedonic contours were used to construct Fisher Ideal price indexes. These indexes appear to have important differences from repeat sales indexes that rely on mean prices that may not control for quality levels. I also found differences between fixed and nonfixed characteristics estimates, and I attributed these to the additional control for the level of characteristics in the fixed *loess* strategy. These differences suggest that in dynamic markets, such as Alameda County, where new housing construction and high levels of remodeling have led to changes in the mean characteristics levels of the housing stock, it is important to control for the confounding effects of price changes and quality changes both in the estimation of the hedonic and in the price index construction. In less dynamic markets, these differences may not be as important.

REFERENCES

Alonso, W. 1964. *Location and Land Use.* Cambridge: Harvard University Press.

Bailey, M., R. Muth, and H. Nourse, H. 1963. “A Regression Method for Real Estate Price Index Construction.” *Journal of the American Statistical Association* 58, pp. 933–942.

Case, K., and R. Shiller. 1989. “The Efficiency of the Market for Single-family Homes.” *American Economic Review* 79, pp. 125–137.

Cleveland, W.S., and S. J. Devlin. 1988. “Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting.” *Journal of the American Statistical Association* 83, pp. 596–610.

_______, _______, and E. Gross. 1988. “Regression by Local Fitting.” *Journal of Econometrics* 37, pp. 87–114.

Diewert, W.E. 1983. “The Theory and Measurement of the Cost-of-living Index and the Measurement of Welfare Change.” In *Price Level*

Measurement, eds. W.E. Diewert and C. Montmarquette, pp. 163–233. Ottawa: Statistics Canada.

_______. 1978. "Superlative Index Numbers and Consistency in Aggregation." *Econometrica* 46, pp. 883–900.

Epple, D. 1987. "Hedonic Price and Implicit Markets: Estimating Demand and Supply Functions." *Journal of Political Economy* 95, pp. 58–80.

Griliches, Z. 1971. "Introduction: Hedonic Price Indexes Revisited." In *Price Indexes and Quality Change,* ed. Z. Griliches. Cambridge: Harvard University Press.

Meese, R., and N. Wallace. 1996. "The Construction of Residential Housing Price Indices: A Comparison of Repeat Sales, Hedonic Regression, and Hybrid Approaches." *Journal of Real Estate Finance and Economics.* Forthcoming.

_______, and _______. 1994. "Testing the Present Value Relation for Housing: Should I Leave My House in San Francisco?" *Journal of Urban Economics* 35, pp. 245–266.

_______, and _______. 1991. "Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices." *Journal of the American Real Estate and Urban Economics Association* 19, pp. 308–332.

Motley, B. 1992. "Index Numbers and the Measurement of Real GDP." Federal Reserve Bank of San Francisco *Economic Review* 1, pp. 3–13.

Muth, R. 1969. *Cities and Housing.* Cambridge: Harvard University Press.

Pollak, R.A. 1991. *The Theory of the Cost-of-living Index.* New York: Oxford University Press.

Quigley, J. 1995. "A Simple Hybrid Model for Estimating Real Estate Price Indexes." *Journal of Housing Economics* 4, pp. 1–12.

Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82, pp. 34–55.

Shiller, R. 1993. *MacroMarkets: Creating Institutions for Managing Society's Largest Economic Risks.* London: Clarendon Press.

Triplett, J.E. 1989. "Price and Technological Change in a Capital Good: A Survey of Research on Computers." In *Technology and Capital Formation,* eds. D.W. Jorgenson and R. Landau, pp. 127–213. Cambridge: MIT Press.

_________. 1987. "Hedonic Functions and Hedonic Indexes." In *The New Palgrave: A Dictionary of Economics,* eds. J. Eatwell, M. Milgate, and P. Newman, pp. 630–634. London: Macmillan.

Wang, F., and P. Zorn. 1995. "Repeat Sales: An Introductory Primer." Unpublished manuscript. Federal Home Loan Mortgage Corporation.