Journal of Econometrics 28 (1985) 307–326. North-Holland

A DYMIMIC MODEL OF HOUSING PRICE DETERMINATION

Robert F. ENGLE, David M. LILIEN and Mark WATSON*

University of California, San Diego, La Jolla, CA 92093, USA
University of California, Irvine, CA 92717, USA
Harvard University, Cambridge, MA 02138, USA

Received November 1983, final version received December 1984

This paper presents and estimates a model of the resale housing market. The data are a cross-section of monthly time series obtained from the multiple-listing service for a suburb of San Diego. The model is specified and estimated as a dynamic multiple indicator multiple cause system of equations where the capitalization rate is taken to be an unobservable time series to be estimated jointly with the unknown parameters. These are estimated by maximum likelihood using an EM algorithm based upon Kalman filtering and smoothing.

The specification of the model features hedonic equations for each house sale and a dynamic equation for the capitalization rate which is constrained to make the expectation of prices equal the present value of the net returns to home ownership whenever the economic variables stabilize at steady state values. Out of steady state, the capitalization rate slowly adapts to new information.

The model attributes a large portion of housing price increases of the 1970's to a fall in the capitalization rate which in turn was driven by rental inflation, tax rates and mortgage rates. Post-sample simulations indicate an initial flattening of housing inflation rates and later a fall brought on by the increase in steady state capitalization rates. In-sample simulations show that although both Proposition 13 and the inflation induced rise in the marginal income tax rates provided partial explanations for the fall in capitalization rates, the single most important factor was the acceleration in price of housing services which interacted with the tax treatment of home ownership to produce an amazing 18% average annual rate of price increase over the last seven years of the 1970's.

1. Introduction

Over the last seven years of the 1970's, the relative price of housing in the United States increased markedly. From January 1973 to January 1980, home ownership costs rose by over 105% while other consumer prices rose by only 78%. In many areas of the country, Southern California in particular, this increase was even more pronounced. In the San Diego suburb we examine in this paper, the average price of a house sold in early 1980 was 275% higher than the average price of a house sold in early 1973.

Neoclassical investment theory suggests that the asset price of a house is simply the capitalized value of its rental services. In the late seventies, much of the increase in housing prices was due to a fall in the rate implicitly used to

*The authors wish to acknowledge support from NSF SOC78-09476 and SOC80-08580 and from the Harvard Graduate Society.

0304-4076/85/$3.30

capitalize current rents rather than any to fundamental change in the demand or supply of housing services. For example, in suburban San Diego, standardized rental prices increased by only 60% between 1973 and 1980, far less than the 275% increase in average house sale price (unstandardized). Of course part of the increase in housing prices may reflect changes in the characteristics of a typical housing unit. But even after controlling for these changes, the price of a suburban San Diego house increased by roughly three times as much as its rental value.

Various explanations have been given for the significant decrease in housing capitalization rates. Over the period in question two factors contributed to a decrease in the effective cost of capital, and thus in capitalization rates. First, the interaction of high inflation rates and the tax system, through the deduction of interest payments, dramatically reduced after-tax real interest rates despite increases in nominal mortgage rates. Second, there was a widespread movement towards reductions in state and local property tax rates beginning with the passage of Proposition 13 in California. In Southern California these reductions lowered property tax rates by approximately 50%.

Many have argued that while these factors have contributed to the rise in housing prices, they are not capable of explaining all of the actual increase. Another explanation that is frequently offered is that of a speculative boom or bubble. In this explanation prices rose merely because they were expected to rise. While such increase cannot be ruled out by economic theory, one would hope that they play a relatively unimportant role in the housing market. If price changes are caused by unexplainable bubbles or waves of optimism rather than more fundamental economic factors, then prices are not functioning properly in their resource allocation role. The increase in housing prices induces capital to flow into the housing market when, from a welfare standpoint, it is more profitably employed elsewhere [see Blanchard and Watson (1982)].

In this paper we attempt to sort out the factors leading to the rise in housing prices by analyzing a high quality micro data set on 915 individual house sales over an 80-month period from a relatively homogeneous suburb of San Diego. We construct a model of both the housing service and asset markets which allows us to measure the empirical importance of the factors discussed above. Differences in rental values across housing units during the same period are attributed to characteristics of the units much as in Grether and Mieszkowski (1974, 1980), or Ridker and Henning (1967) and others. Variations in sale prices over time are explained by asset market behavior and in particular by the capitalization of future rents and costs.

Because our data set is essentially a time series of cross-sections with an important dynamic character, the model has important statistical aspects. The basic capitalization rate each month is treated as an unobserved component which is a determinant of the individual sales price, and which is in turn determined by a set of factors, such as tax rates, mortgage rates, and the market clearing rental rate. Because of the fundamentally dynamic nature of the relationships, the model is called a Dynamic Multiple Indicator Multiple Cause model, a direct generalization of the MIMIC model of Goldberger and Jöreskog (1977). The statistical techniques for this model were developed in Engle and Watson (1981) and Watson and Engle (1983). This paper, the first application of the DYMIMIC method to an empirical problem, demonstrates the power of this formulation for economic research and the ability of the computer software to analyze a problem of this magnitude.

The paper is organized as follows. In section 2 we present the model, and section 3 briefly discusses the estimation technique. The data are discussed in section 4, and in section 5 we present the estimation results as well as the results of specification tests. In section 6 we evaluate the forecast performance of the model for 1980 and 1981, unquestionably the most difficult forecast period for housing in the post-war period. In section 7 we carry out some dynamic simulations to measure the relative importance of various factors in the housing price explosion of the 1970's. Our results suggest that one need not appeal to bubbles to explain the price explosion. Rather the cause of the increase is well explained by market fundamentals. The most important cause of the price increases in the market was rental inflation which can be attributed to demographic shifts. The interaction of inflation, nominal interest rates, and tax policy is also found to be important. Less important was the effect of property tax changes brought about by Proposition 13.

2. The model

Housing is a hedonic good and the rental price of any individual unit must take into account this heterogeneity. The estimation of hedonic indices is a well-studied procedure initiated by Griliches (1961) and introduced into housing economics by Ridker and Henning (1967). In more recent contributions such as Butler (1980), Linneman (1980), Kain and Quigley (1975), Grether and Mieszkowski (1974, 1980), Straszheim (1973, 1975), and many others, various issues in functional form, integration over markets, and estimation methods have been discusssed. The theoretical underpinnings of the model are described by Rosen (1974) who shows that the hedonic index is a synthesis of the demand and supply of heterogeneous goods where the index maps out the equilibrium set of attribute prices which clear the market.

Few of these studies have observations at more than one point of time. Grether and Mieszkowski adjust for the fact that their data are not all at the same point by including an estimated time trend assumed to be the increase in land value. Ferguson and Wheaton (1980) explicitly estimate models of rates of change, but use only new construction and subtract out the rise in the land value. Their measures are consequently more oriented toward construction costs than the prices or rentals of existing housing.

Letting r_{it} be the log of the rental value of a unit i in time period t, and x_{it} be the vector of characteristics of this unit, a typical semi-log specification of the hedonic index in one year would be

$$r_{it} = x_{it}\beta_t + \varepsilon_{it}. \tag{1}$$

When observed over several time periods presumably the index shifts upward due to shifts in the demand and possibly supply schedules underlying the index. A simple specification which allows such a shift employs a rental index r_t^0 which is the log of the rent of a standardized unit with fixed characteristics x^0. The dynamic version of the hedonic index could be written as

$$r_{it} = r_t^0 + (x_{it} - x^0)\beta + \varepsilon_{it}. \tag{2}$$

This specification assumes that the percentage markup of the rental price of a particular unit over the standard unit remains constant over the period. This is a testable restriction which, in this data set, can be accepted as is discussed in the empirical section.

Because the sample of units used in this study are owner occupied houses, the implicit rentals r_{it} are not observed. Instead, data on the log selling prices, p_{it} are used in the model. The relationship between the asset and rental price of a unit is the capitalization rate whose log will be denoted θ_t. Thus

$$p_{it} = r_{it} - \theta_t. \tag{3}$$

Several studies [see Linneman (1980) and Butler (1980)] have estimated hedonic indices over both rental and owner occupied units by treating the capitalization rate as a fixed but unknown constant. While this may be appropriate in a cross-section, it seems unattractive in a time series. Economic models of the capitalization rate emphasize its dependence on expectations of future inflation, tax rates and interest rates. Indeed the discussion in the introduction suggests that a major part of the increase in housing prices can be attributed to decreases in the capitalization rate, rather than to rentals which more closely represent the tightness of the market for housing services. The same capitalization rate should apply to the standardized housing unit implying

$$p_t^0 = r_t^0 - \theta_t. \tag{4}$$

Although both θ_t and p_t^0 are unobservable, r_t^0 is observed (possibly with error) as the BLS index of rental rates in San Diego.

To complete the model an equation is required for the capitalization rate. In a perfect foresight, zero transaction cost housing market, such an equation could be derived from the arbitrage condition that equates the price an investor or home owner would be willing to pay for a house to the presented discounted flow of net rental services from the house. In fact, the housing market satisfies few of the assumptions necessary for a perfect asset market. Houses are heterogeneous. Information is imperfect and costly to acquire. Transaction costs (commissions, moving costs, etc.) can be quite large. Any asset price behavior based on standard arbitrage conditions cannot be expected to hold at each moment of time. Rather these conditions are useful only in the constraints that they place on the long-run properties of the model. The approach taken in this paper is to treat the actual capitalization rate as a random variable that is constrained to follow an arbitrage condition in steady state while allowing a wide range of possible adjustment paths in the short run.

Let Θ^* be the steady state capitalization rate which depends upon steady state rental inflation rates π^*, mortgage rates (which also represent the pre-tax opportunity cost of capital) i^*, and marginal income and property tax rates τ_y^* and τ_p^*, respectively. Depreciation μ is defined as the percent of the asset value which is required to maintain the quality of the unit constant over time. The cost of capital is taken to be $i^*(1-\tau_y^*)=c^*$ although risk adjustment may also be appropriate. The costs of owning a unit at time t with value P_t can therefore be expressed as $\delta P_t = (\mu + \tau_p^*(1-\tau_y^*))P_t$ where it is implicitly assumed that property taxes are a fixed proportion of property values. This assumption is appropriate in pre-Proposition 13 California, but not strictly so subsequently. However, if resale is a consideration, then the market value of the unit will reflect taxation at the new assessed value, hence the formulation may be appropriate in this regime as well.

Given these definitions and assuming that in steady state there are no excess profits to be made, the price an investor would offer would just equate the discounted value of net revenues from the unit. In the steady state, the variables need no time subscripts and the horizon can be taken at infinity,

$$P_t = \sum_{j=0}^{\infty} \frac{R_{t+j}}{(1+c^*)^j} - \sum_{j=0}^{\infty} \frac{\delta P_{t+j}}{(1+c^*)^j}, \tag{5}$$

where R_t and P_t are the rents and values at time t. Substituting the steady state relationships $R_{t+1}=(1+\pi^*)R_t$, $\Theta^* P_t = R_t$, (5) can be rewritten as

$$1 = [\Theta^* - \delta] \sum_{j=0}^{\infty} \left(\frac{1+\pi^*}{1+c^*}\right)^j. \tag{6}$$

Solving for $\theta^* = \log \Theta^*$ gives

$$\begin{aligned}\theta^* &= \log\left(c^* - \pi^* + \delta(1 + c^*)\right) - \log(1 + c^*) \\ &\cong \log\left[i^*\left(1 - \tau_y^*\right) - \pi^* + \left(\mu + \tau_p^*\left(1 - \tau_y^*\right)\right)\left(1 + i^*\left(1 - \tau_y^*\right)\right)\right] \\ &\quad - i^*\left(1 - \tau_y^*\right). \end{aligned} \tag{7}$$

This expression differs only slightly from others such as Muth (1982) or the simple version in Summers (1981).

Eq. (7) defines the steady state value of the capitalization rate as a function of the steady state values of interest rates, tax rates, and the rate of rental inflation. Any sensible model of the capitalization rate should imply this type of long-run behavior. The short-run behavior of the rate will depend on a variety of things including the way in which expectations of interest rates, rental values, and tax policy are formed. Transactions costs and the information structure of the market will also be important determinants of the short-run behavior of the capitalization rate. Short-run economic models emphasize these features to differing degrees depending on the focus of the model. While there is general agreement concerning the steady state properties of the capitalization rate there is no strong concensus concerning its short-run properties. We will take an agnostic and data based approach.

First we construct a short-run approximation to the steady state capitalization rate which we denote by θ_t^*. This variable is constructed by replacing the starred variables appearing in (7) with smoothed moving averages of their observed values. In steady state these moving averages become the starred values so that θ_t^* mimics the long-run behavior of θ^*. Next we allow the actual capitalization rate to vary around θ_t^* in a flexible but stationary manner. The deviations of the rate, θ_t, from θ_t^* are brought about by the observed stationary variables $(z - z^*)$ and an unobserved stationary disturbance term. The actual specification that we employ is

$$\theta_t - \theta_{t-1} = (1 - \phi)\left(\theta_t^* - \theta_{t-1}\right) + (z - z^*)_t\gamma + \xi_t. \tag{8}$$

Here z^* denotes the steady state value of z. In steady state, therefore, θ_t will converge to θ^* as z converges to z^* and the transient dynamics die out. Clearly more general specifications with the same long-run dynamics are possible. Our specification is parsimonious, and as we show in section 5, is not rejected by the data.

Combining eqs. (1)–(4), (7) and (8) gives the model

$$p_{it} = r_t^0 - \theta_t + x_{it}\beta + \varepsilon_{it}, \tag{9}$$

$$\theta_t = \phi\theta_{t-1} + (1 - \phi)\theta_t^* + (z - z^*)_t\gamma + \xi_t, \tag{10}$$

where $-x_0\beta$ has been included in the constant term of eq. (9). The unknown parameters are β, ϕ, γ, μ and the variances.

3. The estimation method

The model presented above is a special case of the dynamic multiple-indicator multiple-cause (DYMIMIC) model discussed in Engle and Watson (1981). At the heart of the DYMIMIC model is a vector of unobserved factors which evolve over time. Part of the evolution of the factors is described by unobservable stochastic disturbances. These unobservable factors are used to describe part of the process generating a vector of observed variables, which are called 'indicators' of the factors. The process generating the indicators may also include exogenous variables and disturbance terms.

In the model presented in the last section the capitalization rate, θ_t, is unobserved. Its evolution is described by eq. (10), the causal equation, and θ_t^* and $z_t - z_t^*$ are its observable causes. The transaction prices, p_{it}, serve as indicators of the capitalization rate, and the set of equations in (9) are called indicator equations.

The stochastic structure of the disturbance terms must be specified before estimation is discussed. We will assume that all disturbances are independently normally distributed with mean zero, that the disturbances in the indicator equations have constant variance σ_ε^2, and that the disturbances in the causal equation have constant variance q. This assumption implies that ξ_t is white noise and uncorrelated with disturbances in the indicator equation.

With these assumptions maximum likelihood estimation is reasonably straightforward. The most obvious approach would be to recursively substitute the causal equation into the indicator equations. The resulting reduced form would express each transaction price as a function of the rental index, house characteristics, a complicated distributed lag of the causal variables, and a complicated disturbance term. The likelihood function could then be formed and maximized with respect to the unknown parameters.

An easier method is available. Suppose, for a moment, that the capitalization rate were observed. Maximum likelihood estimates could then easily be calculated by forming the appropriate sample moment matrices necessary for the multivariate regression problem [e.g., Theil (1971, ch. 7)]. Conversely, if the parameters of the model were known, then standard signal extraction techniques, as described in Whittle (1963) or Anderson and Moore (1979), could be used to estimate the capitalization rate and its variance. These could then be used to form the sample moment matrices of the data. Putting these two procedures together we have an algorithm. From an initial guess of the parameters we use a Kalman filter and smoother, a signal extraction procedure, to estimate the capitalization rate and its variance. Combining these with the observed data we form the appropriate moment matrices and obtain new parameter estimates using standard regression formulae. These new parameter estimates are used to form new estimates of the capitalization rate, and the procedure is repeated until convergence.

The algorithm is a special case of the EM algorithm described in Dempster, Laird and Rubin (1977). Their results show that the final parameter estimates satisfy the first-order conditions for maximization of the likelihood function. Details of the algorithm for the DYMIMIC model can be found in Watson and Engle (1983).

When the final parameter estimates are obtained, the information matrix can be formed using the expression derived in Engle and Watson (1981). The smoothed estimates of the capitalization rate are also produced as a byproduct of the estimation procedure. Conditional on the parameter estimates these are minimum mean square error estimates.

4. The data

Housing prices and specific characteristic data are from single-family homes resold in University City, a suburb of San Diego. They were compiled from multiple listings supplied by the San Diego Board of Realtors. The multiple-listing data are unique in that they include most of the characteristics of a house that may be of interest to a potential buyer, as well as sale price, sale date, and market time. A menu of variables available from the multiple listings is shown in table 1.

Data were collected on a monthly basis from July 1973 to March 1980. These are close-of-escrow dates and we make the assumption that the transaction occurred one month prior to this data. After editing the data, 915 observations were available.[1] Unfortunately, some of the older multiple listings were missing and, therefore, we are missing data for five months, December 1974, February 1975 and October through December 1977. In addition, we had data on only one transaction in January 1975. Average selling price over the period is shown in fig. 1.

The rental data were constructed from the rental component of the San Diego Consumer Price Index. This was published quarterly until 1978 and every two months thereafter. We interpolated the data to get a monthly series.[2]

[1]A referee has pointed out that our results may be subject to a sample selection bias. It could be argued that 'lemons' – houses with undesirable unmeasured characteristics – are placed on the market more frequently than 'non-lemons'. This suggests that our sample of 915 housing sales includes a disproportionate number of 'lemons'. To the extent that these unmeasured characteristics are correlated with x, our coefficients in the indicator equation will be biased. If this bias is important then our model must be viewed as a description of the dynamic behavior of the market price of houses (i.e., prices of housing that appeared on the market) rather than the value of housing (i.e., the implicit price of all housing).

[2]More elaborate interpolation methods [e.g., Ansley and Kohn (1983), or Harvey and Pierse (1984)] could have been used. In section 5 we present the results of a test for measurement error in our rental index. Our failure to reject the null hypothesis of no measurement error suggests that our simple interpolation scheme was adequate for our purposes.

Table 1

Menu of variables from multiple listings.

Financial Variables	Dummy Variables
Listing price	Pool
Selling price	Wall-to-wall carpeting
Type of loan acquired by buyer	Drapes
Type and amount of loan of seller	Patio
Property taxes before sale	Covered patio
Real estate commission	Enclosed patio
	Laundry room
Continuous Variables	Laundry area
Number of bedrooms	220V to property
Number of baths	Sprinklers
Age	Paved street
Lot size	Yard access for boat or RV
Square feet of living area	Dishwasher
Room sizes	View
Distance to bus	Type of construction
Distance to shopping	Zoning
	Type of garage
	Type of heating
	Type of air conditioning
	Type of TV antenna
	Type of sewer service
	Type of fencing
	Type of built-ins
	Condition – excellent, good, fair or poor
	Fireplace
	Type of flooring

State and federal marginal tax rates were supplied by the Internal Revenue Service and the California Franchise Tax Board. We calculated the tax rates, based on average family income and family size, taken from a model of the San Diego economy maintained by Criterion West. Finally, property tax rates were supplied by the San Diego Assessor's Office, and mortgage rates were from the Federal Home Loan Bank Board's survey of mortgage rates on 80% commercial loans for the Los Angeles area.

The use of a market-wide rental index can be criticized on several grounds. Rents in University City may be different from market-wide rents. Furthermore, most, if not all of our data are from owner occupied housing, and the value of housing services in owner occupied units may be different from the value in non-owner occupied housing. Both of these suggest errors-in-variables problems. We can, however, test for these problems. Note that a measurement error in the rental index at time t will be common in all of the indicator equations at time t. This implies that there are two common unobservables in the indicator equations: the capitalization rate and the rental index measure-

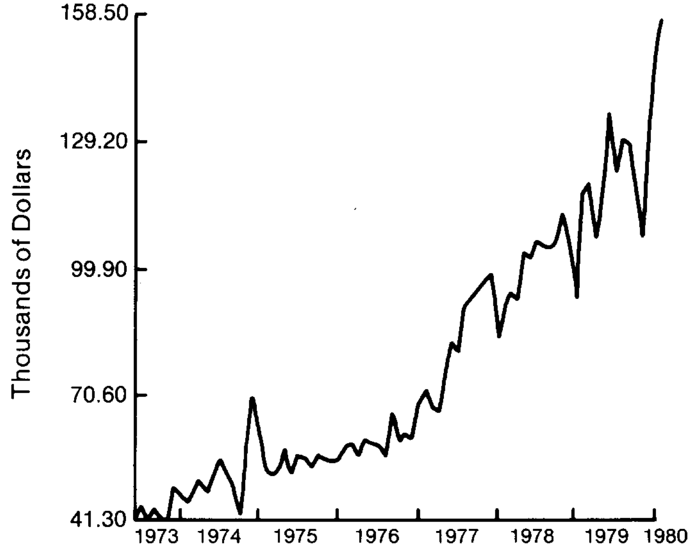

Fig. 1. Average selling price.

ment error. Testing for the presence of a second unobservable is a test for measurement error. This test is carried out in the next section.

The variables entering the underlying capitalization rate were formed as follows:

$$\pi_t^* = (1/12) \sum_{i=0}^{5} a_i \pi_{t-i},$$

$$i_t^* = (1/12) \sum_{i=0}^{5} a_i i_{t-i},$$

where $a_i = (6 - i)/21$, so that the weights decline in a linear fashion, π_t is the annual percent change in the rental index, and i_t is the annual mortgage rate. The variables τ_p^* and τ_y^* were formed as centered 12-month moving averages of the tax rates described above.

5. Estimation results

After some experimentation with the variables presented in table 1, the following characteristics were chosen for use in the indicator equation:

BA Bathrooms per square foot,

AGE Age of house in years,

Table 2

Estimates of the causal equations.[a]

	Model 1	Model 2	Model 3
θ_{t-1}	0.963 (0.034)	0.957 (0.040)	0.955 (0.038)
θ_t^*	0.037	0.043	0.045
$i_{t-1} - i_{t-1}^*$		−0.180 (1.029)	
$i_{t-2} - i_{t-2}^*$		0.207 (1.054)	0.057 (0.490)
$\pi_{t-1} - \pi_{t-1}^*$		0.267 (0.963)	
$\pi_{t-2} - \pi_{t-2}^*$		−1.342 (0.942)	−1.112 (0.509)
$\mu \cdot 10^3$	6.81 (1.92)	7.79 (2.11)	7.85 (1.96)
$q \cdot 10^4$	2.10 (0.93)	2.07 (0.92)	2.04 (0.92)
L	1832.34	1834.93	1834.35

[a] Standard errors are in parentheses.

AGE2	Age squared,
NOTX	Condition of house was not excellent,
LOT	Lot size in square feet,
NR	Number of rooms,
NR2	Number of rooms squared,
ARS	Average room size in square feet,
ARS2	Average room size squared,
3CG	Dummy variable for three-car garage,
SHT	Dummy variable for shake or tile roof,
VU	Dummy variable for view,
CVG	Dummy variable for converted garage,
VIN	Dummy variable for vintage which takes on the value 1, if house was built after 1970,
PL	Dummy variable for pool.

Since all of the sample data were chosen from the same neighborhood, we were able to disregard any neighborhood characteristics which are typically used in similar studies, e.g., Grether and Mieszkowski (1974, 1980) or Ridker and Henning (1967).

The model presented in section 2 assumed that the coefficients on these individual house characteristics were constant over the sample period. Before estimating the complete DYMIMIC model this assumption was tested using a fixed effects model. This model assumes that the capitalization is a fixed unknown constant for each time period, and is easily estimated using a set of

Table 3

Model 3 indicator equations.

Variable	Coefficient[a]		Scale
CONST	−0.850	(0.210)	
BA	0.729	(0.168)	$\times 10^{2}$
AGE	−0.121	(0.031)	$\times 10^{-1}$
AGE2	0.361	(0.131)	$\times 10^{-3}$
NOTX	−0.693	(0.108)	$\times 10^{-1}$
LOT	0.783	(0.198)	$\times 10^{-5}$
LOT2	−0.132	(0.062)	$\times 10^{-9}$
NR	0.334	(0.025)	
NR2	−0.192	(0.023)	$\times 10^{-1}$
ARS	0.404	(0.048)	$\times 10^{-2}$
ARS2	−0.296	(0.067)	$\times 10^{-5}$
3CG	0.672	(0.157)	$\times 10^{-1}$
SHT	0.410	(0.063)	$\times 10^{-1}$
VU	0.884	(0.621)	$\times 10^{-2}$
CVG	−0.989	(0.187)	$\times 10^{-1}$
VINT	−0.367	(0.102)	$\times 10^{-1}$
POO	0.587	(0.091)	$\times 10^{-1}$
σ_ε^2	0.021	(0.030)	$\times 10^{-2}$

[a]Standard errors are in parentheses.

time specific dummy variables. The indicator equations were estimated over the entire sample period; the sample was then split and the usual Chow test carried out. The resulting *F*-statistic was 0.62 which can be compared to the 5% critical value of 1.65. The assumption of constant coefficients appears to be a reasonable one. The complete results for this fixed effect model are not reported here, but can be found in Engle, Lilien and Watson (1981).

The results for various specifications of the DYMIMIC causal equations are reported in table 2, and table 3 presents the estimated indicator equations for model 3. (The results for other models are very similar and are not reported.) The results seem sensible. The capitalization rate adjusts reasonably slowly to a change in the underlying rate. For model 1 half of the adjustment occurs in eighteen months, while for model 3 the figure is fifteen months. Increases in mortgage rates, apart from their effect through θ_t^*, tend to increase the capitalization rate, but this effect is small and statistically insignificant. On the other hand, models 2 and 3 suggest that increases in inflation have a negative and (in model 3) a statistically significant effect on the capitalization rate. The point estimates of the maintenance and depreciation rate correspond to annual rates between 8% (model 1) and 9.5% (model 3). These appear to be rather high, but their standard errors suggest that these estimates are far from precise. Surprisingly, they are close to the figures allowed by the 1981 Tax Reform Act.

Turning now to the indicator equation presented in table 3, all of the coefficients are significantly different from zero, with the exception of the coefficient on *VU*, which has a '*t*-statistic' of 1.4. The fit is reasonably good;

the estimated standard deviation of ε_{it} is approximately 8%. The quadratic terms indicate that there are decreasing returns to house size, lot size, and number of rooms. The negative coefficient for the coefficient on *CVG* indicates that a room from a converted garage is worth less than a standard room. The negative sign on *VIN* indicates that people prefer older homes, once depreciation is taken into account. Both *SHT* and *3CG* are measures of the quality of the home and their coefficients are both significantly positive. The estimates imply that a pool adds about 6% to the value of a home. This result seems at odds with assessment practices. Assessors typically add very little to their estimated value of a home for a pool.

Using model 3 as the null hypothesis some of the assumptions underlying the model were tested. All of the models presented in table 2 impose the constraint that the coefficients on θ_{t-1} and θ_t^* sum to one. Relaxing this constraint produced very similar point estimates and a likelihood ratio statistic of 1.30, well below the critical value for a X_1^2, random variable.

Model 3 was also estimated allowing for a second unobservable factor, assumed to follow AR(1) process. This is a very general specification test for the model under consideration. As mentioned in the previous section this second unobservable could represent [AR(1)] measurement error in our rental index. The absence of the second unobservable implies that measurement error is not a significant problem. A test for a second unobservable is also a direct test for the presence of a 'speculative bubble'. A discussion of bubbles and associated empirical tests can be found in Flood and Garber (1980), Blanchard and Watson (1982), and Burmeister and Wall (1983). In this interpretation the capitalization rate and rental index determine the 'fundamental price' and the second unobservable determines the bubble component. If a rational bubble is present we expect the AR coefficient to be larger than one. The absence of the second unobservable implies the absence of a rational speculative bubble. (Burmeister and Wall carry out a test much like this in their investigation of bubbles and the German hyperinflation.) Finally this test is quite useful in detecting misspecified dynamics in the equation for the capitalization rate. Any misspecification in that equation should leave some serial correlation in the data that the second unobservable will 'mop up'. The likelihood ratio statistic for the presence of the second factor is 1.02.

A few technical points should be made about this test statistic. Two additional parameters were estimated in the model allowing measurement error: the autoregressive parameter and the variance of the disturbance term. Testing for the absence of a second unobservable involves testing whether the variance of its disturbance terms is zero. The fact that the variance is non-negative suggests that a one-tailed test should be used. If the autoregressive parameter were known the correct test would be the square root of the likelihood ratio statistic which would be distributed as a standard normal random variable truncated from below at zero. [See, for example, Gourieroux, Holly and Monfort (1980).] Indeed, under the null this would be the asymptotic distribution of the square root of the likelihood ratio statistic for any arbitrary value of the autoregressive parameter. As Davies (1977) has shown the square of the likelihood statistic, with an estimated autoregressive parameter, is distributed under the null, as the maximum of a sequence of truncated standard normal random variables. [This problem is discussed in detail in Watson and Engle (1984).] Clearly, the critical value for the maximum of a set of standard normal variables is greater than the critical value for a single standard normal. Since $(1.02)^{\frac{1}{2}}$ is below the critical value for a standard normal, we find no significant evidence in favor of a second unobserved component. This finding adds further support for our model. It implies that there is no significant evidence of measurement error or of misspecified dynamics in the equation for the capitalization rate. It also implies that there is no significant evidence for the presence of a speculative bubble. This rejection of the presence of a speculative bubble is of particular importance since it implies that prices are performing their proper role as signals in the competitive market.

Fig. 2 shows our estimates of monthly capitalization rates, Θ_t $(=\exp\theta_t)$ and Θ_t^* $(=\exp\theta_t^*)$, calculated from model 3. Θ_t decreased substantially over the sample period, decreasing from 1.34% to 0.63%. Our rental index, on the other

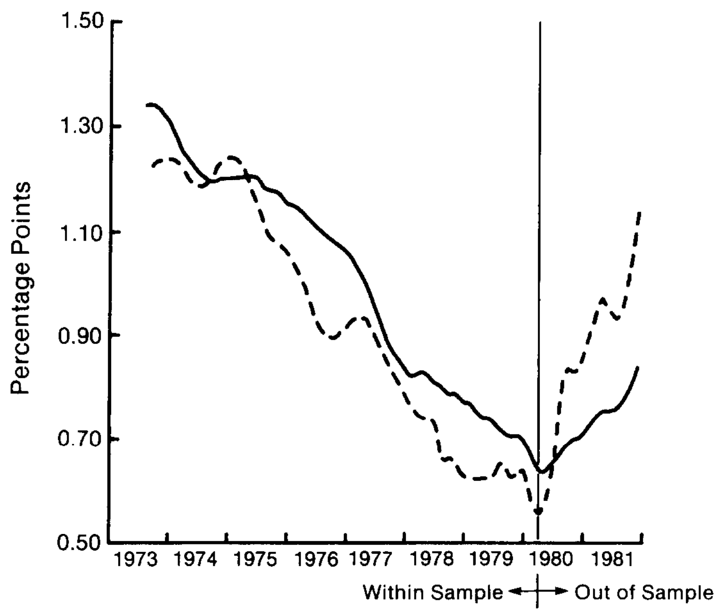

Fig. 2. Capitalization rate.

hand, increased only 60% over the sample period. Combining the capitalization rate and the rental index we can compute an index of housing values. This index increased 235% over the sample period. Clearly, a large fraction of this increase was due to the fall in the capitalization rate.

6. Forecast performance

The end of the sample period was March 1980 and occurred just as mortgage interest rates broke through 13% on their way to over 18%. Such a substantial rise in the cost of funds would normally be expected to have a disasterous effect on the housing market. If these rates are viewed as temporary as opposed to 'steady state', however, then sellers may refuse to reduce their asking prices and the main effect will be a decrease in transactions. It is exactly this difference between steady state capitalization rates and short-run changes which are modeled by the causal equation (10). Thus it is of interest to see whether, faced with the independent variables from 1980 to 1981, the model predicts a substantial fall in property values.

The model was simulated for 16 months ending in July 1981. Over the year and a half from January 1980 to July 1981, rental rates increased at an annual rate of 7% while interest rates rose from 13% to 16.75%, an increase of 29%. The steady state capitalization rate increased over 60% from 0.57% to 0.93%. However, the estimated capitalization rate increased far less, to only 15% above the January figure and the estimated price fell only very slightly, to 99%. Within this period, prices peaked in April and then fell back to just below the January level. The Kalman filter also calculates a standard error for the forecast the rate of change of housing prices. This is simply the standard error of the forecast of the log capitalization rate conditional on all the estimated parameters. In July 1981 this was 4.3%. The fact that the steady state capitalization rate was so far below the current rate, suggested that the market would be likely to fall further as the adjustment took place. This was confirmed by a further simulation through November 1981. Faced with further interest rate bad news, the steady state capitalization rate increased to over twice its January 1980 value while prices fell to 91%. The simulated capitalization rates are also plotted in fig. 2.

One must wonder what the actual housing market performance was like over this period. Directly comparable data are not available as the model is predicting price changes of a standardized unit. Furthermore, the extensive use of 'creative financing' whereby the seller holds substantial paper at below market interest rates, means that selling prices overstate the actual values. Nevertheless, average selling prices of all residential units in this area were obtained for this period. As each observation corresponded to a rather small number of sales, symmetric five-month averages were used to obtain more accurate estimates of the relevant selling prices. As before, it was assumed that a sale occurred one month before it was recorded. Taking January 1980 as a base, there was a decline of 1% by January 1981 although it had increased by 7% in the intervening July. By July 1981, there was a net increase of 13% but by November prices had fallen again. The July 1981 figure of 13% increase is compared with the model prediction of a 1% decline, with a 4.3% standard error. While these are significantly different by conventional criteria and assuming correctly measured prices, the model performance seems quite good in switching from a growth rate of over 20% per year over the sample period, to a flat or slightly declining set of prices in the forecast period. It provides a sensible compromise between the sharp fall to be predicted by static economic analysis and a smooth continuation of past trends.

7. Simulation results

The large fall in the capitalization rate can be attributed to a number of changes that occurred over the sample period. An increase in the rate of inflation, progressive nominal marginal tax rates and the deductability of interest and property tax payments worked together to lower the flow cost of homeownership per dollar of asset value. Proposition 13, enacted in 1978, reduced property tax rates by roughly 50% further reducing the flow cost of homeownership. San Diego changed from a slowly growing city (brought about by reduced military spending in the early 70's) to a rapidly growing city (brought about through the growth of 'high-tech' industry). This rapid growth increased the demand for housing services increasing the rate of rental inflation from 2% below the national average at the beginning of the sample to nearly 2% above the national average at the end of the sample. This acceleration in rental value directly increased housing values, while at the same time increasing expected capital gain from homeownership. This expected capital gain lowered the net real cost of homeownership.

To judge the relative importance of these factors we have simulated the causal equation from our model over the sample period using different realizations of the weakly exogenous variables. While we believe that these simulations are useful and suggestive, they should be interpreted with caution. Our causal equation describes the dynamic adjustment of θ to θ^*. There is no reason to believe that this adjustment process will remain unchanged when the processes generating the exogenous variables change. Put another way, our simulations are subject to the well known Lucas critique. Also, we have not considered feedback between the capitalization rate and our weakly exogenous variables. There is every reason to suspect that some feedback exists. A decrease in marginal tax rates, for example, will increase the capitalization rate and decrease house prices. This decrease in prices reduces residential construction, modifying the stock of housing at some future date. This change in the housing stock causes a change in rental values. While our simulations do not

Table 4

Simulation results.

Simulation[a]	$\Theta^*_{73'6}$ [b]	$\Theta^*_{80'3}$ [b]	$\Theta_{80'3}$ [b]	$\dot{p}^*$ [c]	$\dot{p}$ [d]
0	1.2%	0.6%	0.6%	19%	18%
1	1.2%	0.6%	0.6%	19%	17%
2	1.2%	1.2%	1.2%	7%	7%
3	1.2%	0.6%	0.7%	17%	16%
4	1.0%	0.8%	0.8%	10%	9%
5	1.2%	0.7%	0.7%	16%	15%
6	1.3%	0.9%	0.8%	12%	13%
7	1.3%	0.9%	0.9%	12%	13%
8	1.4%	0.9%	0.9%	13%	14%

[a]Simulation explanation:
0 = fitted values of θ_t given all the data,
1 = dynamic simulation of θ_t for base case,
2 = θ^* fixed at $\theta^*_{73'6}$ for entire period,
3 = no Proposition 13,
4 = rental inflation rate set at U.S. rental inflation rate,
5 = marginal tax rate set at 26%,
6 = real interest payment deductions,
7 = marginal tax rate set at 26% with real interest payment deductions,
8 = no interest payment deductions.

[b]Capitalization rates are in monthly rent/price.

[c]Annual rate of inflation over 80 months assuming steady state throughout.

[d]Annual rate of inflation over 80-month sample period assuming initial value is in steady state. For simulation 0, this can be estimated as well, consequently the estimated growth rate appear slightly different.

consider this type of feedback they do give some rough idea of the relative importance of factors which caused changes in the capitalization rate and prices.

Table 4 presents the simulations. All begin in 1973'6 and end in 1980'3 and are therefore within sample simulations. For each simulation several statistics are presented for the initial and final periods. The calculated value of the steady state capitalization rate Θ^* $(= e^{\theta^*})$ and the estimated value of Θ are given. These imply log prices of standardized units in or out of steady state according to the following definitions:

$$p_t = r_t - \hat{\theta}_t, \tag{11}$$

$$p_t^* = r_t - \theta_t^*, \tag{12}$$

and average annual housing inflation rates $\dot{p}$ and $\dot{p}^*$, respectively.

When historical values of all the weakly exogenous variables are used, two different estimates of θ_t in (11) are reported. In simulation 0, the 'smoothed' estimate is repórted which gives the best estimate using all the data on the individual p_{it}, both past and future. The alternative estimate in simulation (1) does not use any information on the p_{it} and is calculated simply by a dynamic simulation of eq. (10). For historical input, one might expect the former to perform much better than the latter because the dependent variables would keep the simulation 'on track' much the way one-period-ahead forecasts are usually much more reliable than full dynamic simulations in standard econometric analysis. In this case, the estimated θ's are quite similar, giving further support to the model's specification. The estimated growth rates of prices do differ slightly due to the inability of the dynamic simulation to estimate the initial price level. For simulations 1–8, these are taken to be same as the steady state and therefore 1 is the base case for comparison.

The remaining simulations are counter-factual and therefore can only be calculated as a dynamic simulation. Simulation 2 fixes θ^* at its first-period value over the entire sample. Any increase in prices is entirely caused by the increase in rental values. Comparing simulations 1 and 2 we see that a large fraction of the actual increase in prices is attributable to the decrease in θ. The effect of Proposition 13 is shown in simulation 3, which fixes the property tax rate for 1978–1980 at its pre-Proposition 13 level. End-of-sample housing prices are 18% lower than in simulation 1, and the steady state value is 33% lower. The average annual steady state growth rate is 17% rather than 19% in the base case.

To measure the impact of demographic shifts the model was simulated allowing the San Diego rental index to increase at the same rate as the U.S. rental index. This has two effects. First, the value of the rental index is different in every period except the first, so that there is a direct effect on prices. Second, rental inflation and expectations of future rental inflation affect expectations of capital gains and hence the capitalization rate. The results are shown in simulation 4. The initial capitalization rate is lower and the initial price is higher because of the relatively higher rate of rental inflation in the U.S. at the beginning of the period. Conversely, the final-period P is considerably lower because of the relatively lower rate of rental inflation at the end of the period in the U.S. This gives an annual average rate of price increase of only 9% or roughly half of the actual rate of price increase.

The interaction of inflation and the tax system is investigated in simulations 5–8. Changes in the rate of inflation in concert with the tax system have two effects on the capitalization rate. First, an increase in inflation together with progressive nominal marginal tax rates leads to an increase in real marginal tax rates. Since interest and property tax payments are tax deductible, this causes a decrease in the flow cost of homeownership (per dollar of asset value) and a decrease in the capitalization rate. Second, since tax laws allow the deduction of nominal interest payments, an increase in the nominal rate of interest brought about by an increase in inflation leads to a decrease in the after-tax real rate of interest. Thus the capitalization rate agains falls.

Simulation 5 looks at the first of these effects. For this simulation, we have fixed marginal tax rates at their 1973 level of 26%. This leads to a small increase in the final period θ and a larger decrease in θ^*. Simulation 6 uses actual tax rates, but allows only the deduction of real interest payments, while simulation 7 fixes marginal tax rates at 26% and allows only the deduction of real interest payments. Finally, simulation 8 presents results assuming no deductions for interest payments. The three simulations limiting interest deductability imply substantially lower rates of housing price increases, regardless of the exact tax assumptions.

These simulations suggest two primary causes of the rapid increase in housing prices over our sample. The first was demographic shifts which significantly accelerated the rate of rental inflation. The second was the interaction of this inflation and the tax system. Proposition 13 appears to have played a less important role.

8. Conclusions

In this paper, we have presented and estimated a model of the resale housing market. The data were a cross-section of time series obtained from the multiple-listing service for a suburb of San Diego. The model was specified and estimated as a dynamic multiple-indicator multiple-cause system of equations where the capitalization rate was taken to be an unobservable time series to be estimated jointly with the unknown parameters. These were estimated by maximum likelihood using an EM algorithm based upon Kalman filtering and smoothing.

The specification of the model featured hedonic equations for each house sale which standardized the units, and a dynamic equation for the capitalization rate. This equation forced the capitalization rate to stabilize at a 'steady state' rate if all the economic variables stabilized at constant levels and rates. The adjustment to this rate was gradual and allowed shocks to the system to only slowly affect the capitalization rate. Presumably, transaction costs and heterogeneous expectations allow this market to function out of steady state but these factors disappear when the economic situation stabilizes.

When the model is used to forecast housing behavior for the first two years of the 1980's, these factors become very important. The high interest rates increase the steady state capitalization rate dramatically, however, the estimated rate rises only slowly and prices are predicted to rise slightly and then decline. These predictions were supported by market experience.

Simulations are employed to sort out the fundamental causes of the housing price inflation assuming the model to be correctly specified. The overall conclusion is that although both Proposition 13 and the inflation induced rise in marginal income tax rates provided partial explanations, the single most important factor was the demographically driven acceleration in the cost of housing services which interacted with the tax treatment of homeownership to produce the amazing 18% average annual rate of increase.

References

Anderson, B.D.O. and J.B. Moore, 1979, Optimal filtering (Prentice Hall, Englewood Cliffs, NJ).

Ansley, C.F. and R. Kohn, 1983, Missing data problems and dynamic regressions, Mimeo. (University of Chicago, Chicago, IL).

Blanchard, O.J. and M.W. Watson, 1982, Bubbles, rational expectations, and financial markets, in: Paul Wachtel, ed., Crises in the economic and financial structure (Lexington Books, Lexington, MA).

Butler, R.V., 1980, Cross-sectional variation in the hedonic relationship for urban housing markets, Journal of Regional Science 20, 439–453.

Davies, R.B., 1977, Hypothesis testing when a nuisance parameter is present only under the alternative, Biometrika 64, 247–254.

Dempster, A.P., N.M. Laird and D.B. Rubin, 1977, Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society B 39, 1–39.

Engle, R.F. and M. Watson, 1981, A one factor multivariate time series model of metropolitan wage rates, Journal of the American Statistical Association 76, 774–780.

Engle, R.F., D.M. Lilien and M.W. Watson, 1981, A dynamic model of housing price determination, Discussion paper 81-2 (University of California, San Diego, CA).

Ferguson, R.F. and W.C. Wheaton, 1980, The determinants of housing inflation in the '70s: Metropolitan housing price indices, Working paper (Joint Center for Urban Studies).

Flood, R.P. and P.M. Garber, 1980, Market fundamental versus price-level bubbles: The first test, Journal of Political Economy 88, 745–770.

Goldberger, A.S., 1977, Maximum-likelihood estimation of regressions containing unobservable independent variables, in: D.J. Aigner and A.S. Goldberger, eds., Latent variables in socio-economic models (Academic Press, New York).

Gourieroux, C., A. Holly and A. Monfort, 1980, Likelihood ratio test, Wald test and Kuhn–Tucker test in linear models with inequality constraints on the regression parameters, Econometrica 50, 63–80.

Grether, D.M. and Mieszkowski, 1974, Determinants of real estate values, Journal of Urban Economics 1, 127–146.

Grether, D.M. and Mieszkowski, 1980, The effects of non-residential land uses on the prices of adjacent housing: Some estimates from proximity effects, Journal of Urban Economics 15, 1–15.

Griliches, Z., 1961, Hedonic price indices for automobiles: An econometric analysis of quality change, in: Models of income determination, Studies in income and wealth, Vol. 28 (Columbia University Press for NBER, New York).

Harvey, A.C. and R.G. Pierse, 1984, Estimating missing observations in economic time series, Journal of the American Statistical Association 79, 125–131.

Kain, J. and J.M. Quigley, 1975, Housing markets and racial discrimination: A microeconomic analysis (Columbia University Press for NBER, New York).

Linneman, P., 1980, Some empirical results on the nature of the hedonic price function for the urban housing market, Journal of Urban Economics 8, 47–68.

Muth, R.F., 1982, Effects of the U.S. tax system on housing prices and consumption, in: Ron Grievson, ed., Conference volume (University of California, Santa Cruz, CA).

Ridker, R.G. and J.A. Henning, 1967, The determinants of residential property values with special reference to air pollution, Review of Economics and Statistics 44, 246–257.

Rosen, S., 1974, Hedonic prices and implicit markets: Product differentiation in pure competition, Journal of Political Economy 82, 34–55.

Straszheim, M.R., 1973, Estimation of the demand for urban housing services from household interview data, Review of Economics and Statistics 55, 1–8.

Straszheim, M.R., 1975, An econometric analysis of the urban housing market (Columbia University Press for NBER, New York).

Summers, L.N., 1981, Inflation, the stock market and owner occupied housing, AER Papers and Proceedings 71, 429–434.

Theil, H., 1971, Principles of econometrics (Wiley, New York).

Watson, M.W. and R.F. Engle, 1983, Alternative algorithms for estimating dynamic factor, MIMIC and varying coefficient regression models, Journal of Econometrics 23, 385–400.

Watson, M.W. and R.F. Engle, 1984, Testing for regression coefficient stability with a stationary AR(1) alternative, Review of Economics and Statistics 67, 341–346.