Evaluating House Price Forecasts

Authors John M. Clapp and Carmelo Giaccotto

Abstract

This study uses an autoregressive process to model a city-wide house price index. The model is used to produce one-quarter ahead forecasts for individual properties. We propose that managers use a battery of tests to compare prediction errors (PEs); in particular, their empirical distribution reveals important information.

Transaction data from Dade County, Florida is used. PEs from two forecasting models, hedonic and repeat sales, show some departure from desirable properties of forecasts. Also, both show some informational inefficiency, but the hedonic is more efficient than the repeat. Nonparametric smoothing shows that the hedonic method dominates the repeat over an important range of PEs; thus, many risk-averse managers might prefer a forecast based on the hedonic method.

Introduction

Real estate decision makers would benefit from accurate forecasts of house prices and of the variance of future prices. For example, developers and investors could decide whether the projected returns from a housing development were sufficient to offset the risks. Prospective homeowners could balance the consumption value of a home against risks and rewards from the investment component. Real estate appraisers, who provide information to mortgage lenders, could decide whether to report that neighborhood house prices are declining, stable or increasing.[1]

The literature on housing market efficiency has demonstrated that house prices exhibit some inertia over short-to-intermediate timeframes. For example, Case and Shiller (1989) found that, contrary to weak-form efficiency, between 25% and 50% of a (real) price index change in one year persisted into the following year. Kuo (1996) improved the Case-Shiller methodology for testing weak-form efficiency by jointly estimating the price index and the serial correlation parameters within the context of the repeat sales model. Kuo assumed a second order autoregressive process to model the rate of change in the price index; his research implies some predictability up to four quarters ahead.

Interestingly, the semi-strong form of the efficient market hypothesis has not fared well either; Case and Shiller (1990), Clapp and Giaccotto (1994) and others, used a number of macro and local economic variables to forecast prices and excess returns to housing for periods up to one year ahead. These variables include local unemployment, expected inflation, mortgage payments, income and population age. This evidence shows that house prices do not quickly impound publicly available information. Moreover, the research of Mankiw and Weil (1989) suggests that prices may be predicted up to twenty years into the future as a result of current and predictable changes in adult population. However, their hypothesis remains controversial (see Hendershott, 1991; and DiPasquale and Wheaton, 1994). These studies, like those in the previous paragraph, are designed to test various kinds of market efficiency.

Another part of the literature deals with methods for forecasting house prices. One branch of this literature uses purely time series data to forecast house prices. Brown, Song and McGillivray (1997) add to earlier studies of British housing by allowing some coefficients of the forecasting equation to vary over time. They use various transformations of the mean percentage error to test model performance. This study employs a similar methodology, especially with the use of Theil's U-Statistic.

Zhou (1997) uses a VAR model with time series data to find that the volume of sales and house prices are cointegrated.[2] He conducted several tests of forecasting power using regressions on predicted values. This study uses these tests on prediction errors (PE's); in addition, a variety of parametric and nonparametric techniques are combined to test for desirable properties of PE's.

Pace, Barry, Gilley and Sirmans (2000) use data on sales of individual houses to construct and forecast prices for a standard house. They use a semiparametric spatial-temporal forecasting model, whereas this study employs a Bayesian parametric method. Their approach is similar to the approach in this study in that both evaluate forecasts by constructing cumulative distribution functions of prediction errors. The two differ in that the focus of this study is on how to evaluate forecasts whereas they focus on forecasting methods.

This study proposes a framework for evaluating the accuracy of alternative house price forecasting models. In the forecasting literature, a number of techniques exist for comparing first and second moments of out-of-sample forecast errors (Theil, 1966; and Puelz and Sobol, 1995). These are reviewed and applied to the data.

Managers can use various criteria when evaluating the accuracy of any forecasting method.[3] The criteria for accurate one-step-ahead forecasts include zero mean prediction errors (PEs) and normally distributed PEs (*i.e.*, well-behaved skewness and kurtosis).[4] The mean squared prediction error should be low relative to total variation in the prices to be predicted. The forecasts should be informationally efficient in the sense that PEs are unrelated to predicted price; moreover, PEs should be unrelated to local-market information (*e.g.*, age of the house) available at the time the prediction was made.

This study compares the repeat sales and hedonic forecasting methods using the above criteria and nonparametric estimates of the entire empirical distribution of

PEs. It argues that a manager can maximize utility by applying utility weights to the PEs in the tails of the distribution, as well as to those in the middle. Graphical nonparametric Probability Density Functions (PDFs) and Cumulative Density Functions (CDFs) are shown to resolve certain conflicts that may arise when using statistics such as the means, skewness and kurtosis of the PEs.

Thus, the contributions of this study may be summarized as follows:

1. The first to thoroughly examine how a manager should evaluate house price forecasts that use sales prices of individual houses.[5]
2. It combines existing parametric and nonparametric tests in a comprehensive battery of tests that will be useful to managers attempting to choose among alternative forecasting methods.
3. It shows that the full nonparametric PDFs of PEs can be compared graphically. Visual comparison allows managers to apply their own utility weights to forecasting errors. This enables the decision maker to resolve certain conflicts that may arise with other evaluation methods. For example, the method with lower mean out-of-sample prediction error may have higher skewness; the graphical method allows a rational basis for choice between the two.

To obtain the house price forecasts used in this study, a generalized version of the repeat sales model originally proposed by Bailey, Muth and Nourse (1963) is employed. An alternative set of forecasts was obtained from the hedonic model of Rosen (1974). In each case, an autoregressive process with a unit root was used to model the time series behavior of house prices. However, the focus of this study is on evaluating the forecasts, not on how they are estimated: Any two forecasting models (*e.g.*, from the flip of a coin and from extrapolation of trend) could be used.

The remainder of this article is organized as follows. In the next section, the local price level is modeled as a stochastic process with inertia. The two competing forecasting methods are proposed. In the following section, a number of preferred properties is outlined for forecast errors, and there is a discussion of the informational efficiency of forecasts using Theil's (1966) decomposition. Next, a random sample of properties that sold from 1976 through the fourth quarter of 1995 in Dade County, Florida is used to estimate the hedonic and repeat sales models. An out-of-sample forecast is then computed for a group of properties not used in estimation and the prediction errors using the statistical methodologies outlined earlier are analyzed. The final section is the conclusion.

Two Methods for Forecasting House Price Changes

Suppose a cross-sectional time series sample of house prices plus a number of house characteristics is observed. If the street address is available, then one may construct a repeat sales index as in Bailey, Muth and Nourse (1963) or Case and

Shiller (1989); alternately, a variation of Rosen's (1974) hedonic model may be estimated. In each of these cases, the price index is typically treated as a fixed quantity to be estimated as the coefficients in a linear regression model. Hence, these methods do not attempt forecasts based on past and present values of the overall price level; they ignore the evidence on the forecastability of house prices reviewed in the preceding section.

The local (*e.g.*, city-wide) price level should be treated as a time series process. Then an estimate of the unknown parameters of the stochastic process allows forecasts of individual house prices one period ahead.

House prices are likely to share time series characteristics with the consumer price index; in fact, housing constitutes about 40% of the CPI. This suggests that the house price level may contain a unit root and the first differences display serial correlation perhaps over one or two years. This hypothesis cannot be tested directly with a unit root test (a Dickey-Fuller test) because the price level is a latent variable, therefore unobservable. Indirect evidence for the presence of a unit root, however, may be found in Kuo (1996) and Case and Shiller (1989). Hence, an autoregressive model with a unit root in the level of the city-wide price index is proposed. Specifically, define y_{it} as the (log) price of the ith house sold at time t. This price may be decomposed into three parts: (1) the city-wide price level; (2) the value of a set of locational and structural characteristics; and (3) a mean zero house specific error term. Case and Shiller (1987), Hill, Sirmans and Knight (1999) and Englund, Gordon and Quigley (1999) assume the second term follows a Gaussian random walk, but treat the city-wide price level as a deterministic constant. This specification is equivalent to a heteroscedastic model where the variance is proportional to time between sales.

The forecasting model starts with the same decomposition of price; however, the emphasis is on the city-wide price level rather than the idiosyncratic noise term. The rate of growth in the price level is treated like a stochastic process with a large predictable component; once the model is estimated, individual house price forecasts may be obtained from the city-wide index.

Thus, define C_t to be the (log) level of the index at time t; the associated first difference (rate of growth) is $\Delta C_t = C_t - C_{t-1}$. This difference is hypothesized to follow an autoregressive process of order p:

$$\Delta C_t = \mu + \phi_1 \Delta C_{t-1} + \cdots + \phi_p \Delta C_{t-p} + \varepsilon_t, \tag{1}$$

Where $t = 0, 1, \ldots, T$ and T represents the total number of time periods. The innovation term ε_t is assumed to behave like white noise with variance $\sigma^2 q_0$. The systematic part of this model $\mu + \phi_1 \Delta C_{t-1} + \cdots + \phi_p \Delta C_{t-p}$ represents a sustainable trend, whereas the error term ε_t has no lasting effect on prices. The autoregressive coefficient ϕ_j represents the percentage of the growth rate in period

$t - j$ that carries over to time t. This model is similar to the one proposed by Kuo (1996). The typical tools of time series analysis (*e.g.*, auto correlation function) are not applicable here because C_t itself is not observed. Hence, determining the order p in Equation (1) is problematic. Kuo arbitrarily set $p = 2$; this study experimented with various values of $p = 2$, 4 and 6.

Equation (1) may be imbedded into one of the traditional models for estimating price indices. For example, suppose that at each point in time t a sample of transactions is observed. Let Y_t be the vector of individual sales prices and Z_t the matrix of corresponding hedonic characteristics. Then, the hedonic model of Rosen (1973) may be written as:

$$Y_t = Z_t\beta_t + e_t, \tag{2}$$

and the vector of hedonic coefficients β_t is assumed to follow a Markov process:

$$\beta_{t+1} = \Phi_{t+1}\beta_t + W_{t+1}. \tag{3}$$

The autoregressive process in Equation (1) may be embedded into Equation (3); the Kalman filter then provides an efficient method for estimation and forecasting (see Harvey and Peters, 1990).

The repeat sales model of Bailey, Muth and Nourse (1963) is obtained from Equation (2) by setting Y_t equal to the (log) price difference between the first and second sale. The corresponding matrix Z_t contains the typical dummy values -1, $0 + 1$ corresponding to the first sale, no sale and second sale, respectively. Forecasts at the individual property level are easily obtained from Equations (2) and (3).

The two forecasting models examined here are not essential to the purposes of this article. Any competing forecasting methods could be used.

Empirical Methods

Empirical methods are designed to evaluate the out-of-sample forecasting accuracy for a cross-section of properties at a given point in time. These tests are conducted for a short time horizon. Thus, the tests focus on the accuracy of forecasts at the individual house level rather than on time series properties of the forecasts.[6]

Preferred Properties of Forecast Errors

Decision makers would like forecast errors to possess a set of preferred distributional properties. At the very least, the manager should expect unbiased forecasts on average. Moreover, although a particular methodology might behave well in the middle of the distribution, it may display a tendency to overpredict (left skewness) or underpredict (right skewness), or make gross errors on both sides of the distribution (kurtosis). Homogeneity of variance is another preferred property for prediction errors.

In sum, the set of preferred distributional properties is: zero mean, constant variance across properties, no skewness or kurtosis; that is, forecast errors (PEs) should be approximately normally distributed and fairly homogeneous. The Gaussian assumption about PEs is not based on the theory of rational expectations. But, it "provides a simple form that captures some desirable features of error weighting," (Gershenfeld and Weigend, 1994:64). In practice, the normality of prediction errors for a given cross-section of properties may follow for this reason: If the prediction errors are caused by many small independent variables, then the central limit theorem implies that they will be asymptotically distributed according to the normal distribution.

These criteria help managers decide whether a forecasting model is an adequate representation of reality. A normal distribution of PE's is one indication that large systematic components of the process generating the data have been included in the model.[7]

To check whether a set of forecast errors display these preferred properties, the first four moments of the empirical distribution are computed. In large samples, the third moment (skewness) has a normal distribution with zero mean and standard deviation of $\sqrt{6/n_t}$; similarly, the fourth moment (kurtosis) is normal with mean of 3 and standard deviation of $\sqrt{24/n_t}$. A non-parametric test is also used for normality of the PEs: the Shapiro-Wilk test (W). The W-Statistic is a good omnibus test based on ordered errors and their variance. Details on this test may be found in Conover (1980: 364–67). Since the PEs may be non-normal in many time periods, Wilcoxon's signed rank test that the median prediction error is equal to zero is also used [*i.e.*, price forecasts are unbiased (see Gibbons and Chakraborti, 1992)].

It is very difficult to check all possible ways that the variance of PE may vary across individual properties; for example, in the hedonic model the variance may be related to size of the property, to building age or to a latent variable. A solution to the problem is to use a variable that encompasses as many of the hedonic characteristics as possible; a good candidate would be the expected house value $E(y_{i\,t+1})$. Hence, the following regression is used to check for constant variance across properties:

$$\hat{e}^2_{i,t+1} = \gamma_0 + \gamma_1 E(y_{i,t+1}) + \gamma_2 E(y_{i,t+1})^2 + \zeta_{i\,t+1}, \tag{4}$$

where $\hat{e}_{i,t}$ is the prediction error and $y_{i,t}$ is the log of house price for property i at time t. If the variance is constant, zero slopes should be found; positive slopes imply greater error variance for more valuable homes. The estimated predicted value $\hat{y}_{i,t+1}$ for $E(y_{i,t+1})$ will be used.

Relative Efficiency of Hedonic vs. Repeat Forecasts

Decision makers are interested in a forecast that has the lowest possible variance of PE's. Unfortunately, in practice this lower bound is unknown; so it cannot be known whether a particular forecasting model is efficient in a statistical sense. Since there are two competing methodologies, the focus is on whether forecasts based on the hedonic model are more or less efficient than repeat sales forecasts.

The degree of efficiency can be judged from Theil's U^2 Statistic, which is based on the Mean Squared Prediction Error (MSPE). Define the MSPE as: $1/n_t \, \Sigma_{i=1}^{n_t} (y_{i,t} - \hat{y}_{i,t})^2 = 1/n_t \, \Sigma_{i=1}^{n_t} \hat{e}_{i,t}^2$. Then, the U^2 Statistic, for each period t, may be defined as:

$$U_t^2 = 1.0 - \sum_{i=1}^{n_t} (y_{i,t} - \hat{y}_{i,t})^2 \Big/ \sum_{i=1}^{n_t} (y_{i,t} - \bar{y}_{i,t})^2. \tag{5}$$

In a similar fashion to the coefficient of determination, R^2, from linear regression, Theil's U^2 Statistic may be used to measure ''goodness of fit'' between the actual and predicted house prices. Values of U^2 close to 1.0 imply highly efficient forecasts in the sense that the variance of the forecast error is nearly equal to the variance of actual prices.

This statistic will be useful in ranking forecast performance of the repeat sales model versus the hedonic model. To actually test the hypothesis that one methodology is more or less efficient than the other, a statistic first proposed by Granger and Newbold (1986) is used. Consider the regression:

$$\hat{e}_{i,t}^{\mathrm{Rpt}} - \hat{e}_{i,t}^{\mathrm{Hed}} = \alpha_0 + \alpha_1(\hat{e}_{i,t}^{\mathrm{Rpt}} + \hat{e}_{i,t}^{\mathrm{Hed}}) + \varepsilon_{i,t}. \tag{6}$$

The ordinary least squares estimator of α_1 is equivalent to:

$$\hat{\alpha}_1 = \frac{\mathrm{Var}(\hat{e}_{i,t}^{\mathrm{Rpt}}) - \mathrm{Var}(\hat{e}_{i,t}^{\mathrm{Hed}})}{Var(\hat{e}_{i,t}^{\mathrm{Rpt}} + \hat{e}_{i,t}^{\mathrm{Hed}})}, \tag{7}$$

large and positive (negative) values of $\hat{\alpha}_1$ are consistent with the hypothesis that the repeat model is less (more) accurate than the hedonic method. The test statistic for this hypothesis is:

$$\rho = \frac{r\sqrt{n_t - 2}}{\sqrt{1 - r^2}} \tag{8}$$

where r is the correlation coefficient between the sum and difference of prediction errors. The statistic ρ will follow a t-distribution with $n_t - 2$ degrees of freedom.

Informational Efficiency of the Forecasts: Theil's Decomposition

A *one step ahead* forecasting methodology that uses all information available at time t can be tested for informational efficiency. An efficient forecast should be uncorrelated with information available at the time the forecast was made. If a significant correlation exists, then a manager may be able to use this contemporaneous information to improve the forecast.

Orthogonality of forecast errors can be tested by checking whether $\mathrm{Cov}(\hat{e}_{t+1,} z_t) = 0$, where z is a property characteristic. Once again there is a problem as to which characteristics one should use; as in the previous section, the expected house price value $E(y_{i\,t+1})$ is used.

$$\hat{e}_{i,t-1} = \gamma_0 + \gamma_1 E(y_{i,t+1}) + \gamma_2 [E(y_{i,t+1})]^2 + \zeta_{i,t+1}, \tag{9}$$

if the intercept and slope parameters are zero, then the forecasting methodology is not throwing away valuable information, and it is not possible to use any of the information available at time t to improve the forecast. The innovation in this study is to require informational efficiency for each of the cross-section of properties that traded for each forecasted time period. Equation (9) can be tested in linear form ($\gamma_2 = 0$) as well as with all parameters unconstrained.

As a further test of information efficiency, whether prediction errors are correlated with hedonic characteristics such as age and assessed value is tested. For the repeat sales model, the covariance between the first sales price and predictions error is examined.

If there is informational inefficiency, then it is desirable to know both its source and its economic importance. The source of inefficiency in the forecasting process may be revealed by decomposing the MSPE into its three main components: degree of bias, variability and unexplained random errors in the forecasting process (Theil, 1966). Consider the regression of actual sales prices on predicted prices during a given period t:

$$y_{i,t} = \alpha_0 + \alpha_1 \hat{y}_{i,t} + u_{i,t}. \tag{10}$$

Let $\bar{\hat{y}}_t$ and $\bar{y}_t$ denote the average predicted and actual price, respectively, for period t. Theil pointed out that the mean squared prediction error can be decomposed as follows:

$$\text{MSPE} = (\bar{\hat{y}}_t - \bar{y}_t)^2 + (1 - \hat{\alpha}_1)^2 \,\text{Var}(\hat{y}) + \text{ESS}/n_t, \tag{11}$$

where $\hat{\alpha}_1$ is the ordinary least squares estimator of α_1 in Equation (10). Var($\hat{y}$) is the cross-sectional variance of the predicted values from the forecasting model for period t and ESS is the error sum of squares from Equation (10).

Divide both sides of Equation (11) by MSPE to decompose MSPE as follows:

$$1 = U^{\text{Bias}} + U^{\text{Regression}} + U^{\text{Error}}. \tag{12}$$

An unbiased, informationally efficient forecasting model would have $\alpha_0 = 0$ and $\alpha_1 = 1$. Since regression Equation (10) passes through the means, $\alpha_0 = 0$ occurs when there is no bias: This corresponds to U^{Bias}, the first term in Equations (11) and (12), being equal to zero. When $\alpha_1 = 1$, cross-sectional variations in $\hat{y}_{i,t}$ are equal to variations in actual price. When this occurs, the second term in Equation (12) (the regression bias term) goes to zero. Thus, when both $\alpha_0 = 0$ and $\alpha_1 = 1$, there is an informationally efficient and unbiased forecast with mean-squared error being composed entirely of the residual from Equation (11). However, it is important to realize that the conditions $\alpha_0 = 0$ and $\alpha_1 = 1$ are just sufficient conditions for unbiased forecasts, they are not necessary. In fact, any combination of these two parameters such that $\alpha_0 - (1 - \alpha_1)E\hat{y}_{i,t} = 0$ is the necessary and sufficient condition for zero bias (Holden and Peel, 1990). Hence, large deviations from the pair (0, 1) may still be consistent with unbiased forecasts.

Comparing the Two Distributions of PEs

Tests in the previous section focused on evaluating only the first and second moments of prediction errors for each of the two methods (repeat and hedonic) separately. But, decision makers may be sensitive to the entire distribution of PEs. A graphical comparison of the CDF's of the PEs can provide the manager with relevant information. Kernel-smoothing methods are used to estimate the CDF's;

this obviates the need to approximate the empirical probability distribution with a theoretical distribution.

The significance of any differences between the two empirical distributions of the PEs can be tested with the Wilcoxon-Mann-Whitney test and the Kruskal-Wallis test [see Mosteller and Rourke (1973) for details].

A chi-square statistic for the null hypothesis that the two distributions have equal median prediction errors is based on the deviation between the observed number of ranks above the median and its expected value. This chi-square statistic is approximated by a squared standard normal deviate.

Data and Analysis of Prediction Errors

Data and Estimation of Forecasting Models

The original database contains all transactions in the Miami MSA (Dade County, Florida) from 1971 through the first half of 1997 as recorded by the Florida Department of Revenue; the total number of records is over 300,000. The data have been screened to eliminate data errors or transactions that appear to be less than arm's length (*e.g.*, a sales price of $1). For each property, the two most recent sales prices, date of sale (year and month), assessed value, age and square feet were collected.[7]

To keep the estimation problem manageable, a random sample of 5,159 houses that sold twice during the interval from the first quarter of 1976 through the second quarter of 1997 were selected. Specifically, for the repeat model, there are 4,372 properties with first sale during the period 1976Q1 through 93Q4 (Q1–Q72), and second sale from 76Q2 through 95Q4 (Q2–Q80). These data are used to estimate the parameters of the forecasting model. The remaining 787 pairs, not included in the estimation sample, have second sale in the forecasting period 1996Q1 thru 97Q2 (Q81–Q86), while the first sale is randomly distributed over the preceding eighty quarters. For the hedonic model, each sale is treated as a one-only sale and the property characteristics are used to control for quality. Thus, the sample size for the estimation period is 9,531 (4,372*2+787), and for the out-of-sample prediction interval, there are 787 observations, exactly the same properties as for the repeat sample.

By the standards of traditional regression models, the size of the database seems relatively small. However, the estimation of the unknown parameters by maximum likelihood requires substantial computer time (even on a large mainframe computer). The subroutines KALMN were used in conjunction with BCONF from the IMSL stat/math library to maximize the likelihood function (subject to boundary constraints).

Analysis of Prediction Errors

The first question of interest is whether the PEs are white noise: zero mean, no skewness or kurtosis and homogeneous variance. Exhibit 1 displays the first four sample moments for each of the last six quarters and all six quarters together. The evidence appears to say that, on average, the repeat method does a slightly better job of forecasting than the hedonic model. The average forecast error ranges from -0.027 to 0.024 for the repeat sample, and from -0.030 to 0.031 for the hedonic; for all quarters taken together, the average error is 0.5% for the repeat, and 0.9% for the hedonic.

The standard deviation for all time periods implies that, under a traditional t-test for zero mean, both hedonic and repeat mean prediction errors are not significantly different from zero (*i.e.*, they are unbiased). However, the estimated kurtosis (which is discussed below) is not consistent with normality, therefore the t-test will not be robust and conclusions based on it may be suspect.

The hedonic model appears to generate negatively skewed forecast errors. The repeat model errors are generally symmetric: over the entire forecast sample period, the coefficient of skewness is only -0.044 with a p-value of 0.614 for the repeat model. However, both models display heavy tails relative to the normal distribution; the sample values are, roughly, 6.0 for the repeat and 8.2 for the hedonic. For a normal random variable these numbers should be much closer to 3.0. Similarly, the p-values for the Shapiro-Wilk W-Statistic show significant departures from normality for both methods; the hedonic model shows larger departures in most subperiods. However, in all quarters, the significance of these results appear to be driven by a smaller standard deviation for the distribution of hedonic PEs. Therefore, conclusions about the relative merits of the two methods depend on further tests.

Efficient use of information by the two forecasting methods can be evaluated by regressing prediction errors on predicted price [Equation (9)]. The first five columns of Exhibit 2 report these results. From a statistical point of view, both methods display informational inefficiency over the entire forecasting timeframe: the estimated value of γ_1 is -0.15 for the repeat model and 0.043 for the hedonic. Both are significantly different from zero, but the repeat method has greater departures from the null hypothesis of no relationship: The repeat method has higher R^2 and F-values. The hedonic regression relationship is statistically significant in only two of the six subperiods (quarters 81 and 86), whereas the repeat regression is statistically significant in all subperiods. Therefore, these tests reveal a greater degree of informational inefficiency for the repeat method. Note with interest that, for this method, the estimated γ_1 coefficient [see Equation (9)] is always negative; this suggests that the repeat model tends to under-predict at the high-end of house prices. Interestingly, just the opposite holds for the hedonic model.[8]

Exhibit 1 | Descriptive Statistics for Prediction Errors (PEs) by Quarter. Hedonic and Repeat Sales Models

Quarter	Number of Observations	Model	Mean	Std. Dev.	Skewness		Kurtosis		Shapiro-Wilk *W* Test	
81	111	Repeat	−0.005	0.157	0.723	(0.002)	4.169	(0.012)	0.959	(0.012)
		Hedonic	−0.030	0.140	−1.269	(0.000)	7.843	(0.000)	0.925	(0.000)
82	172	Repeat	0.023	0.208	0.135	(0.470)	4.524	(0.000)	0.969	(0.029)
		Hedonic	0.024	0.165	−0.926	(0.000)	5.439	(0.000)	0.946	(0.000)
83	165	Repeat	0.017	0.176	0.838	(0.000)	6.258	(0.000)	0.967	(0.017)
		Hedonic	0.014	0.157	−1.292	(0.000)	9.847	(0.000)	0.946	(0.000)
84	141	Repeat	−0.022	0.201	−0.046	(0.824)	3.971	(0.019)	0.973	(0.130)
		Hedonic	−0.012	0.155	−2.311	(0.000)	14.360	(0.000)	0.871	(0.000)
85	116	Repeat	0.024	0.211	0.059	(0.795)	5.729	(0.000)	0.968	(0.066)
		Hedonic	0.031	0.151	−1.197	(0.000)	8.406	(0.000)	0.945	(0.000)
86	82	Repeat	−0.027	0.232	−1.467	(0.000)	8.223	(0.000)	0.904	(0.000)
		Hedonic	0.026	0.164	−0.499	(0.065)	3.470	(0.385)	0.973	(0.272)
All	787	Repeat	0.005	0.198	−0.044	(0.614)	5.969	(0.000)	0.969	(0.000)
		Hedonic	0.009	0.157	−1.209	(0.000)	8.245	(0.000)	0.938	(0.000)

Notes: Quarters 81 thru 86 correspond to 1996Q1 thru 1997Q2. All refers to all six quarters combined. Prediction Errors (PEs) are as a % of predicted price for all properties sold within the indicated forecasting time period. Numbers in parenthesis are *p*-values for the statistics. For normal random variables skewness should be zero and kurtosis 3. Small *p*-values imply that the sample estimate differs from the population moment for a normal variable. The Shapiro-Wilk W-Statistic is a test for normality: *p*-value less than 0.05 indicates rejection of normality at the 5% level.

Exhibit 2 | Informational Efficiency of PEs: Repeat Compaerd to Hedonic

		$\hat{e}_{i,t+1} = \gamma_0 + \gamma_1 E(\gamma_{i,t+1}) + \zeta_{i,t+1}$					$\hat{e}_{i,t+1} = \gamma_0 + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \zeta_{i,t+1}$					
Quarter	Model	γ_0	γ_1	R^2	F-Value	Prob>F	γ_0	γ_1	γ_2	R^2	F-test	Prob>F
81	Repeat	1.08	−0.09	0.04	5.85	0.02	0.88	−0.08		0.05	5.34	0.02
		(2.41)	(−2.42)				(2.30)	(−2.31)				
81	Hedonic	−1.29	0.11	0.05	7.07	0.01	−1.35	0.11	0.04	0.10	5.68	0.00
		(−2.72)	(2.66)				(−3.30)	(3.06)	(1.87)			
82	Repeat	1.78	−0.15	0.07	13.70	0.00	1.03	−0.09		0.04	7.84	0.01
		(3.75)	(−3.70)				(2.86)	(−2.80)				
82	Hedonic	−0.67	0.06	0.01	2.29	0.13	−1.52	0.10	0.11	0.14	14.26	0.00
		(−1.46)	(1.52)				(−3.80)	(3.19)	(5.0)			
83	Repeat	1.58	−0.14	0.08	14.98	0.00	1.14	−0.10		0.07	11.97	0.00
		(3.91)	(−3.87)				(3.51)	(−3.46)				
83	Hedonic	0.01	0.00	−0.01	0.00	0.99	−0.61	0.03	0.09	0.09	7.60	0.01
		(0.43)	(0.04)				(−1.64)	(0.94)	(3.89)			
84	Repeat	1.16	−0.101	0.03	4.62	0.03	0.64	−0.06		0.02	2.31	0.13
		(2.11)	(−2.15)				(1.47)	(−1.52)				
84	Hedonic	−0.45	0.04	−0.00	0.90	0.35	−0.75	0.05	0.04	0.03	2.43	0.12
		(−0.97)	(0.94)				(−1.79)	(1.51)	(1.94)			
85	Repeat	2.49	−0.21	0.14	19.10	0.00	1.80	−0.16		0.12	15.13	0.00
		(4.41)	(−4.37)				(3.93)	(−3.89)				
85	Hedonic	0.12	−0.01	−0.01	0.03	0.86	−0.08	0.00	0.03	0.01	0.46	(0.50)
		(0.51)	(−0.17)				(−0.18)	(0.07)	(0.95)			
86	Repeat	2.34	−0.2	0.14	14.60	0.00	1.15	−0.10		0.05	3.80	0.05
		(3.77)	(−3.82)				(1.90)	(−1.95)				

Exhibit 2 | (continued)

Informational Efficiency of PEs: Repeat Compaerd to Hedonic

		$\hat{e}_{i,t+1} = \gamma_0 + \gamma_1 E(\gamma_{i,t+1}) + \zeta_{i,t+1}$					$\hat{e}_{i,t+1} = \gamma_0 + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \zeta_{i,t+1}$					
Quarter	Model	γ_0	γ_1	R^2	F-Value	Prob>F	γ_0	γ_1	γ_2	R^2	F-test	Prob>F
86	Hedonic	−1.33 (−2.22)	0.11 (2.26)	0.05	5.13	0.03	−1.42 (−2.70)	0.11 (2.57)	0.05 (1.56)	0.09	3.97	0.05
All	Repeat	1.75 (8.50)	−0.15 (−8.48)	0.08	71.92	0.00	1.12 (6.61)	−0.10 (−6.59)		0.05	43.43	0.00
	Hedonic	−0.50 (−2.50)	0.04 (2.54)	0.01	6.47	0.01	−0.84 (−5.01)	0.06 (4.07)	0.06 (6.45)	0.06	25.66	0.00

Notes: Quarters 81 thru 86 correspond to 1996Q1 thru 1997Q2. All refers to all six quarters combined. The F-values are for a one tail test that all parameters (except for the intercept) are zero. In the regression: $\hat{e}_{i,t+1} = \gamma_0 + \gamma_1 E(y_{i,t+1}) + \zeta_{i,t+1}$, the predicted value of each property is used as a proxy for the expected value. In the regression: $\hat{e}_{i,t+1} = \gamma_0 + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \zeta_{i,t+1}$, X_1 is the first sales price for repeats. For the hedonic regressions, X_1 is assessed value and X_2 is building age. Numbers in parenthesis are t-Statistics.

The last five columns of Exhibit 2 evaluate the efficiency of the two forecasting methods by determining whether information available at the time of the forecast is correlated with the prediction errors. For the repeat method, prediction errors are significantly and negatively related to the first price of the repeat pair. For the hedonic method, the assessed value and age of the house are significantly related to the PEs over the whole period and in three of six subperiods. These results are consistent with the regressions of PEs on predicted price; the hedonic method appears to be slightly more efficient than the repeat.

In spite of the statistical significance of the results in Exhibit 2, however, one suspects that the absolute values of the regression coefficients are too small to be of economic value (*e.g.*, to set up a trading rule to earn abnormal returns from trading in the housing market). Theil's decomposition of the mean square prediction errors is used to evaluate the degree and economic significance of efficiency.

The first five columns of Exhibit 3 report the results of the regression of actual price on predicted price, Equation (10). The *t*-Statistics in the first two columns evaluate the null hypotheses that (individually) $\alpha_0 = 0$ and $\alpha_1 = 1$; columns four and five evaluate that joint hypothesis. The estimated values of α_0 and α_1, for the most part, differ significantly from the pair (0, 1) for the entire forecast period as well as most quarterly subperiods. The only exception occurs for the hedonic model in quarters 83, 84 and 85, where the actual values are (0.008, 1.000) in quarter 83, (−0.450, 1.040) in 84 and (0.120, 0.990) in quarter 85.

Note that the rejection of the null hypothesis is driven more by the intercept being different from zero than α_1 different from unity. This result suggests that prediction errors may be biased. For additional evidence on this point, Exhibit 3 displays the estimated value of Wilcoxon's signed rank test. The result suggests that the mean prediction errors for the hedonic method are significantly different from zero in all subperiods, except quarter 84, as well as over the entire forecasting horizon. The repeat mean PEs are never significantly different from zero; the smallest *p*-value of the signed rank test is 0.065 in quarter 84. Therefore, forecasts based on the hedonic model display more bias than those based on the repeat sales method.

Theil's decomposition of this regression (see the last three columns of Exhibit 3) shows that the hedonic method performs well over all quarters: 99% of the mean squared prediction error (MSPE) is due to the regression residual U^{Error}, whereas only 1% is due to bias and regression bias combined. This compares to the repeat method, where 92% is due to U^{Error}, and 8% is due to $U^{\text{Regression}}$. The hedonic method has four subperiods with U^{Error} greater than 95%, whereas the repeat method has only one subperiod where this is true. Thus, departures from desirable properties have greater economic significance in the case of the repeat model.

To analyze further the second order moments of prediction errors, Exhibit 4 reports Theil's U^2 Statistic, a statistic similar to R^2 that evaluates the relative efficiency of the two forecasts. The U^2 Statistic is consistently higher for the hedonic model;

Exhibit 3 | Theil's Decomposition of Mean Square Prediction Errors (MSPE)

		$y_{i,t+1} = \alpha_0 + \alpha_1 \hat{y}_{i,t+1} + u_{i,t+1}$						Theil's Decomposition of MSPE		
Quarter	Model	α_0	α_1	R^2	F-test for $\alpha_0 = 0$ and $\alpha_1 = 1$	Prob>F	Wilcoxon Test	U^{Bias}	$U^{Regression}$	U^{Error}
81	Repeat	1.076 (2.405)	0.91 (2.42)	0.840	6.090	0.02	−1.183 (0.118)	0.001	0.051	0.948
81	Hedonic	−1.291 (−2.72)	1.11 (2.66)	0.870	12.430	0.00	−1.886 (0.030)	0.043	0.058	0.899
82	Repeat	1.781 (3.75)	0.85 (3.70)	0.720	16.150	0.00	−1.099 (0.136)	0.012	0.074	0.914
82	Hedonic	−0.670 (−1.46)	1.06 (1.52)	0.810	5.870	0.02	−3.002 (0.001)	0.020	0.013	0.970
83	Repeat	1.590 (3.91)	0.87 (3.87)	0.790	16.840	0.00	−0.880 (0.189)	0.009	0.083	0.907
83	Hedonic	0.008 (0.019)	1.00 (0.01)	0.820	1.220	0.27	−1.786 (0.037)	0.007	0.000	0.990
84	Repeat	1.160 (2.11)	0.90 (2.15)	0.730	6.350	0.01	−1.515 (0.065)	0.011	0.032	0.957
84	Hedonic	−0.450 (−0.97)	1.04 (0.95)	0.830	1.780	0.18	−0.262 (0.397)	0.006	0.006	0.990
85	Repeat	2.490 (4.41)	0.79 (4.37)	0.700	21.150	0.00	−0.802 (0.211)	0.013	0.142	0.850
85	Hedonic	0.120 (0.24)	0.99 (0.18)	0.820	4.870	0.03	−2.904 (0.002)	0.040	0.000	0.960
86	Repeat	2.340 (3.77)	0.80 (3.82)	0.740	16.300	0.00	−0.594 (0.276)	0.014	0.152	0.834

Exhibit 3 | (continued)

Theil's Decomposition of Mean Square Prediction Errors (MSPE)

		$y_{i,t+1} = \alpha_0 + \alpha_1 \hat{y}_{i,t+1} + u_{i,t+1}$							Theil's Decomposition of MSPE		
Quarter	Model	α_0	α_1	R^2	F-test for $\alpha_0 = 0$ and $\alpha_1 = 1$	Prob>F	Wilcoxon Test		U^{Bias}	$U^{Regression}$	U^{Error}
86	Hedonic	−1.330 (−2.22)	1.12 (2.26)	0.860	7.460	0.01	−1.801	(0.036)	0.025	0.059	0.916
All	Repeat	1.750 (8.49)	0.85 (8.47)	0.750	72.400	0.00	−0.053	(0.479)	0.001	0.084	0.920
	Hedonic	−0.498 (−2.51)	1.04 (2.56)	0.830	9.110	0.00	−3.480	(0.000)	0.003	0.008	0.990

Notes: Quarters 81 thru 86 correspond to 1996Q1 thru 1997Q2. All refers to all six quarters combined. Numbers in parentheses under the column headings α_0 and α_1 are t-Statistics. The Wilcoxon Rank Sign test is for the null hypothesis of zero mean PE during the indicated forecasting time period. Numbers in parenthesis following the Wilcoxon test are p-values. Small p-values indicate a non-zero.

Exhibit 4 | Analysis of Second Order Moments: Theil's U^2 Statistic and Tests for Constant Variance Across Properties

$$\hat{e}^2_{i,t+1} = \gamma_0 + \gamma_1 E(y_{i,t+1}) + \gamma_2[E(y_{i,t+1})]^2 + \zeta_{i,t+1}$$

Quarter	Model	U^2 Statistic	r	γ_0	γ_1	γ_2	R^2
All	Repeat	0.730	0.25 (7.27)	18.926 (2.000)	−3.218 (−9.424)	0.137 (9.398)	0.101
All	Hedonic	0.860		6.965 (2.617)	−1.171 (−2.567)	0.049 (2.526)	0.015
81	Repeat	0.690	0.12 (1.30)	0.921 (0.33)	−0.140 (−0.293)	0.005 (0.264)	0.000
81	Hedonic	0.800		5.070 (0.812)	−0.824 (−0.770)	0.033 (0.732)	0.043
82	Repeat	0.770	0.25 (3.43)	6.912 (1.477)	−1.142 (0.156)	0.047 (0.168)	0.024
82	Hedonic	0.820		1.308 (0.263)	−0.219 (−0.256)	0.009 (0.254)	0.000
83	Repeat	0.710	0.12 (1.58)	7.417 (2.050)	−1.255 (−2.025)	0.053 (2.008)	0.015
83	Hedonic	0.830		19.765 (3.154)	−3.347 (−3.121)	0.142 (3.091)	0.066
84	Repeat	0.650	0.31 (3.94)	15.931 (3.170)	−2.697 (−3.142)	0.114 (3.120)	0.061
84	Hedonic	0.810		−5.173 (−0.590)	0.904 (0.603)	−0.039 (−0.612)	0.000
85	Repeat	0.690	0.35 (4.03)	22.280 (3.535)	−3.723 (−3.459)	−0.156 (3.387)	0.158
85	Hedonic	0.850		14.638 (2.438)	−2.486 (−2.411)	0.106 (2.388)	0.049
86	Repeat	0.830	0.37 (3.58)	53.465 (8.623)	−9.140 (−8.722)	0.390 (8.826)	0.535
86	Hedonic	0.860		8.859 (1.820)	−1.514 (−1.815)	0.065 (1.815)	0.016

Notes: Quarters 81 thru 86 correspond to 1996Q1 thru 1997Q2. All refers to all six quarters combined. Prediction Errors (PEs) are as a % of predicted price for all properties sold within the indicated forecasting time period. Numbers in parentheses are *t*-Statistics. r = simple correlation coefficient from Equation (5), the regression of the difference between the two errors on their sum. In all regressions, the hedonic is more efficient than the repeat ($r > 0$); it is significantly more efficient in four of six periods and in all periods. In the last four columns, the predicted value of each property is used as a proxy for the expected value.

it ranges from 0.80 to 0.86. The range for the repeat sales model is from 0.65 to 0.83. This indicates that the hedonic model explains a larger percentage of the variance in forecasted sales prices than the repeat method.

The simple correlation coefficient, r, is from the regression of the difference of PEs (equal to the repeat PE minus the hedonic PE) on the sum of PEs. It is always positive, indicating that the repeat PE has a larger variance than the hedonic PE. The t-Statistic [ρ, Equation (8)] indicates that the hedonic is significantly more efficient than the repeat in all forecasting quarters and in four out of six subperiods.

Exhibit 4 reports substantial evidence of heteroscedasticity for the two methods. The repeat PEs display more evidence of variability in variance: over all quarters, the repeat coefficients, and their t-values, are larger in absolute value. For the repeat method, significant heteroscedasticity is evident in each quarter from 1983–86. The hedonic method shows significant heteroscedasticity only in quarters 83 and 85. For both methods, the statistically significant results indicate that higher predicted prices are associated with smaller error variance, over the range of the explanatory variable.

The evidence in Exhibit 4 says that local market indicators, as measured by property characteristics, have some power to predict the future variances of house prices. This implies that house characteristics must be among omitted variables that could be used to improve forecasting power; the use of these characteristics could produce PEs that are closer to the normal distribution.[9] For repeat sales, this house-specific information is limited to the first price.

The hedonic method displays less heteroscedasticity than the repeat method: more detailed housing characteristics can help to move the forecasts towards homogeneous variance. Since neighborhoods tend to have homogeneous housing characteristics, this implies that local information adds significantly to forecasts of the housing market, in the sense of removing variables that have systematic effects.

It is straightforward to use the results in Exhibit 4 to re-estimate the forecasts with a correction for heteroscedasticity.[10] More importantly, Exhibits 1 and 4 indicate a conflict between first and second moments: The first moment favors the repeat method whereas the second favors the hedonic. Thus, there is a need to look beyond just first and second order moments.

Managers Benefit from Comparing the Two Distributions of Prediction Errors

Exhibit 5 shows the empirical distribution functions for the prediction errors from the two methods.[11] This gives more specificity to the higher standard deviation and lower skewness and kurtosis observed in Exhibit 1 for the repeat method. Exhibit 5 reveals that the PEs from the repeat method are shifted to the left relative

Exhibit 5 | Repeat and Hedonic PEs

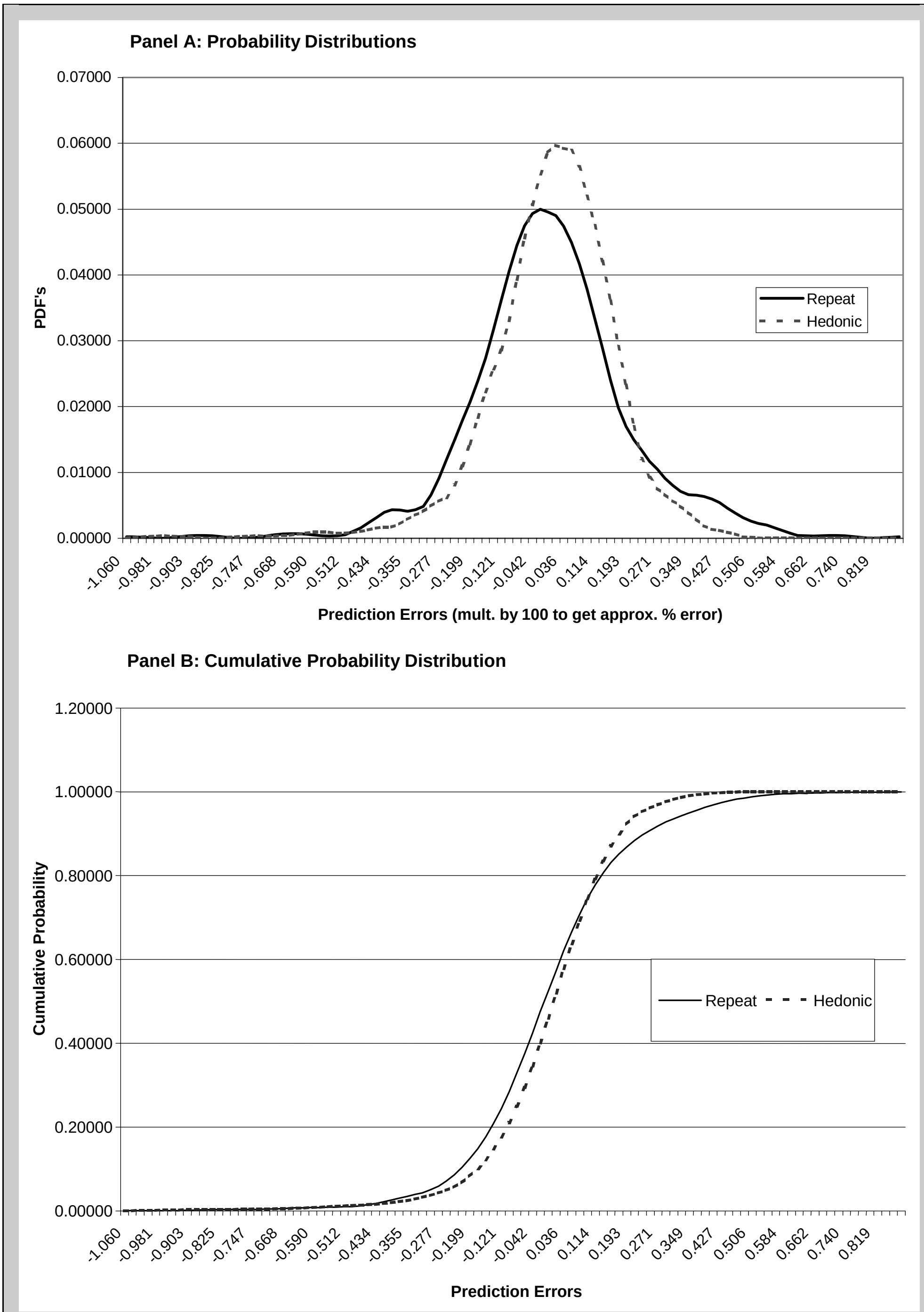

Exhibit 6 | Tests on the Differences Between Repeat and Hedonic Distributions

		Wilcoxon Rank Sums Test			Kruskal-Wallis Test		Points > Median		
Quarter	Observations	$W+$, repeat	Z	Prob>|Z|	Chi-Square	Prob>CHISQ	Sum	Z	Prob>|Z|
81	111	12,505	0.267	0.789	0.721	0.788	55	−0.133	0.894
82	172	28,633	−1.124	0.261	1.266	0.261	78	−1.720	0.085
83	165	26,944	−0.419	0.068	0.176	0.675	80	−0.550	0.583
84	141	19,054	−1.310	0.190	1.720	0.190	65	−1.310	0.191
85	116	12,930	−1.143	0.253	1.310	0.253	52	−1.570	0.116
86	82	6,292	−1.560	0.120	2.430	0.120	39	−0.623	0.533
All	787	600,149	−2.175	0.030	4.730	0.030	364	−2.970	0.003

Notes: Quarters 81 thru 86 correspond to 1996Q1 thru 1997Q2. All refers to all six quarters combined. Prediction Errors (PEs) are as a % of predicted price for all properties sold within the indicated forecasting time period. Test statistics are reported for the absolute values of repeat sales prediction errors; these are ranked, after combining with hedonic absolute PEs. Negative z-values, and chi-square > 1, indicate that the repeat distribution places more weight on negative PEs than does the hedonic distribution.

to those from the hedonic method. For PEs from −0.06 through +0.21, the hedonic distribution dominates (*i.e.*, for every outcome there is a higher probability than in the repeat distribution). Furthermore, a large part of the probability is in this range: 60% of the hedonic probability and 46% of the repeat probability. Thus, use of the hedonic method would give the decision maker 14% greater probability of being within this range.

Panel B of Exhibit 5 shows that the two CDF's cross only once. If the mean hedonic PE were closer to zero (or the same as) the repeat mean, then the hedonic method might dominate the repeat method in the sense that any risk averse decision maker would prefer the hedonic forecast.[12] But, the mean hedonic PE is greater than the mean repeat PE.

A case can be made that a broad class of decision makers would prefer the hedonic forecasts. Consider a developer or an investor who has decided to commit to a new development based on a price forecasted by one of the two methods. When this speculative development is finished, the actual sales price will differ from the forecasted price. With the hedonic method, 60% of the probability mass is concentrated in the range from −6% to +21%. This means that there is a good chance that any error will be neutral to positive (*i.e.*, actual sales price will be above predicted by between zero and 21%, or that the error will be slightly negative). The positive errors imply a pleasant surprise in the sense that the developer and the investors will make more money than they expected.

The repeat method has much more probability in the lower tail. In particular, there is much greater probability of large, negative surprises, ranging up to −50%. Since these decision makers are likely to attach a great deal of negative utility to actual price being substantially below forecasted price, they will maximize expected utility by using the hedonic method.

Exhibit 6 presents statistical tests on the significance of the differences between the two CDF's. The Wilcoxon and Kruskal-Wallis tests show that the distribution of repeat PEs is significantly lower than the corresponding hedonic distribution. That is, the repeat distribution has significantly more density at lower values of the PEs. Furthermore, the binomial tests show that the median of the repeat distribution is significantly lower than the median of the hedonic distribution.[13]

The behavior in the tails of the two distributions is best understood with Exhibit 5. The different kurtosis and skewness measures in Exhibit 1 would not lead one to suspect that the repeat PEs have so much probability mass at lower PE values compared to the hedonic method. The tests in Exhibit 6 show that these differences are statistically significant.

Conclusion

This study focuses on a variety of empirical tests of the accuracy of alternative models for forecasting house prices. The methods proposed here allow managers to choose the best forecasting method (*i.e.*, the one that performs best out-of-sample on a variety of criteria). These criteria are desirable for any risk averse manager.

The hedonic and repeat sales forecasting models based on Equation (1) are estimated with housing transactions from Miami, Florida. Both sets of prediction errors (PEs) show significant departures from the set of desirable properties. The repeat sales method performs better than the hedonic in terms of some basic, descriptive statistics: repeat PEs have lower means, skewness and kurtosis. But, the repeat PEs have larger standard deviations in all forecasted quarters.

The decision criteria proposed here are capable of reversing conclusions based on simple descriptive statistics. For the Miami data, both forecasting methods show some informational inefficiency, but the hedonic is more efficient than the repeat. For example, the hedonic PEs are not as closely related to property characteristics available at the time the forecast was made. Also, the hedonic characteristics explain a lower percentage of the variance of the PEs. Most importantly, the hedonic performed significantly better than the repeat method in terms of Theil's (1966) decomposition of the mean squared prediction error: overall, this decomposition shows that the hedonic method is useful management forecasting tool (Exhibit 3, last column).

Kernel smoothing methods are used to estimate probability distribution functions (and CDFs) for the two distributions of PEs (see Exhibit 5). The hedonic method dominates the repeat over an important range of PEs: from about −6% to +23%. This implies neutral to pleasant surprises for investors, lenders and other decision makers. Furthermore, the repeat method has much higher probability of a large, negative error; this would not be suspected from the skewness and kurtosis numbers in Exhibit 1. Therefore, a case can be made that many risk-averse managers would prefer a forecast based on the hedonic method.

Exhibit 5 provides an essential part of a framework enabling managers to choose among alternative forecasting methods. In theory, this always gives the correct decision because it allows decision makers to apply their own utility weights to outcomes; thus, it resolves conflicts among traditional evaluation criteria such as mean, skewness and kurtosis.

Endnotes

1. The uniform FHLMC residential appraisal form 2055 (revised 11/97) requires the appraiser to check a box for one of these forecasts.
2. A related study by Dua, Miller and Smyth (1999) uses a Bayesian VAR model to forecast aggregate U.S. home sales.
3. This is an essential part of the strategy of evaluating the properties of the forecasts themselves; market efficiency was not tested.
4. See Gershenfeld and Weigen (1994) and Diebold and Lopez (1996).

[5] The problem here is that standard time series methods (*e.g.*, the partial autocorrelation function and/or unit root tests) can not be applied because the data are not evenly spaced over time.

[6] As pointed by Diebold and Lopez (1996), there is no reason to expect serially correlated forecasts when the time horizon of the forecast is one period into the future; that observation, plus the short forecasting period (6 quarters), obviate the need for serial correlation tests.

[7] We use several tests that do not require normal PEs. So normality is a preferred condition, not a required condition.

[8] The authors thank Dean Gatzlaff, Florida State University, for providing the data. A more complete description of the database may be found in Gatzlaff and Ling (1994).

[9] When Equation (9) was estimated with the square term included, the results displayed substantial collinearity. The substantive findings were essentially the same as those reported in Exhibit 2.

[10] Exhibit 2 shows that property characteristics can be used to improve the *level* of the house price forecasts.

[11] However, reworking the model to remove bias is a difficult job, beyond the scope of this article. Most importantly, fitting the model to the forecasting period is avoided.

[12] These empirical CDFs were estimated using standard kernel density estimation techniques.

[13] This discussion is motivated by the concept of second order stochastic dominance (SSD) as applied to security returns (see Ingersoll, 1987). But, that theory does not apply here; higher prediction errors (either positive or negative) subtract from utility, whereas higher returns add to utility. However, the decision maker will want to maximize expected utility over the entire distribution of PEs, regardless of the form of the distribution. Thus, the SSD concept is relevant.

[14] Exhibit 1 showed that the mean repeat PE is lower, but statistical tests based on a normal distribution failed to reveal statistical significance.

References

Bailey, M. J., R. F. Muth and H. O. Nourse, A Regression Method for Real Estate Price Index Construction, *Journal of the American Statistical Association*, 1963, 58, 933–42.

Brown, J. P., H. Song and A. McGillivray, Forecasting U.K. House Prices: A Time Varying Coefficient Approach, *Economic Modeling*, 1997, 14, 529–48.

Case, K. E. and R. J. Shiller, The Efficiency of the Market for Single Family Homes, *American Economic Review,* 1989, 79:1, 125–37.

Case, K. E. and R. J. Shiller, Forecasting Prices and Excess Returns in the Housing Market, *Journal of the American Real Estate and Urban Economics Association*, 1990, 18:3, 253–73.

Clapp, J. M. and C. Giaccotto, The Influence of Economic Variables on Local House Price Dynamics, *Journal of Urban Economics,* 1994, 36, 161–83.

——., Residential Hedonic Models: A Rational Expectations Approach to Age Effects, *Journal of Urban Economics,* 1998, 44, 415–37.

Conover, W. J., *Practical Non-Parametric Statistics*, Second edition, New York, NY: John Wiley, 1980.

Diebold, F. X. and J. Lopez, Forecast Evaluation and Combination, *Statistical Methods in Finance*, In G. S. Maddala and C. R. Rao (Eds.), New York, NY: Elsevier, 1996.

DiPasquale, D. and W. Wheaton, Housing Market Dynamics and the Future of Housing Prices, *Journal of Urban Economics*, 1994, 35:1, 1–27.

Dua, P., S. Miller and D. Smyth, Using Leading Indicators to Forecast U.S. Home Sales in a Bayesian Vector Autoregressive Framework, *Journal of Real Estate Finance and Economics*, 1999, 18:2, 191–205.

Englund, P, T. Gordon and J. M. Quigley, The Valuation of Real Capital: A Random Walk Down Kungsgatan, *Journal of Housing Economics,* 1999, 8, 205–16.

Gatzlaff, D. and D. Ling, Measuring Changes in Local House Prices: An Empirical Investigation of Alternative Methodologies, *Journal of Urban Economics*, 1994, 35:2, 221–44.

Gershenfeld, N. and A. Weigend, The Future of Time Series: Learning and Understanding, *Sciences of Complexity,* Proc. Vol. XV 1–70, Addison-Wesley, 1993.

Gibbons, J. D. and S. Chakraborti, *Nonparametric Statistical Inference*, Third edition, New York, NY: Marcel Dekker, Inc., 1992.

Granger, C. W. J. and P. Newbold, *Forecasting Economic Time Series*, Second edition, New York, NY: Academic Press, 1986.

Harvey, A. C. and S. Peters, Estimation Procedures for Structural Time Series Model, *Journal of Forecasting*, 1990, 9, 89–108.

Hendershott, P. H., Are Real House Prices Likely to Decline by 47 Percent?, *Regional Science and Urban Economics*, 1991, 21:4, 553–63.

Hill, R.C., C. F. Sirmans and J. R. Knight, A Random Walk Down Main Street?, *Regional Science and Urban Economics*, 1999, 29:1, 89–103.

Holden, K. and D. Peel, On Testing for Unbiasedness and Efficiency of Forecasts, *Manchester School of Economic and Social Studies*, 1990, 58, 120–27.

Ingersoll, J. E., *Theory of Financial Decision Making*, Totowa, NJ: Rowman & Littlefield, 1987.

Kuo, C., Serial Correlation and Seasonality in the Real Estate Market, *Journal of Real Estate Finance and Economics*, 1996, 12, 139–62.

Lindgren, B. W., *Statistical Theory*, Third edition, New York, NY: Macmillan, 1976.

Mankiw, N. G. and D. Weil, The Baby Boom, The Baby Bust and the Housing Market, *Regional Science and Urban Economics*, 1989, 19, 235–58.

Mosteller, F. and R. Rourke, *Sturdy Statistics: Nonparametric and Order Statistics*, Redding, MA: Addison-Wesley, 1973.

Pace, R. K., R. Barry, O. Gilley and C. F. Sirmans, A Method for Spatial-Temporal Forecasting With an Application to Real Estate Prices, *International Journal of Forecasting*, 2000, 16, 229–46.

Puelz, A. V. and M. Sobol, The Accuracy of Cross-Validation Results in Forecasting, *Decision Sciences*, 1995, 26:6, 803–19.

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

Shiller, R. J., Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks, Oxford: Clarendon Press, 1993.

Theil, H., *Applied Economic Forecasting*, Rand McNally & Company, 1966.

Zhou, Z., Forecasting Sales and Price for Exiting Single-Family Homes: A VAR Model with Error Correction, *Journal of Real Estate Research*, 1997, 14:1/2, 155–67.

The authors thank the Center for Real Estate and Urban Economic Studies (*CREUES*), *University of Connecticut for financial support, and Dean Gatzlaff* (*Florida State University*) *for providing the data.*

John M. Clapp, University of Connecticut, Storrs, CT 06269-2041 or johnc@sba.uconn.edu.

Carmelo Giaccotto, University of Connecticut, Storrs, CT 06269-2041 or carmelo@sba.uconn.edu.