HOUSING PRICES FORECASTS: AN EVALUATION OF FORECASTING MODELS

Prakash L. Dheeriya, Ph. D, California State University-Dominguez Hills, 1000 E. Victoria Street, Carson, CA 90747, U.S.A., 310-243-3350, pdheeriya@csudh.edu

ABSTRACT

This study deals with the relative performance of econometric models used in forecasting house prices. To evaluate the efficiency of forecasts using these four models, we use the following three yardsticks: root mean squared error, mean absolute error and the Theil U Statistic. Our results indicate that simple exponential smoothing models outperform all other exponential smoothing models as well as the multivariate regression models in all forecasting periods (short, medium-term and long term). The multivariate regression models performed poorly in forecasts of house prices in all periods. We do find a strong bias (E^m) contributed to a major part of the root mean square error in our forecasts.

INTRODUCTION

Many market participants including real estate investors, banks, nonbank financial institutions, portfolio managers, are interested in coming up with a model which accurately predicts local housing prices. Managers of banks, REITS, homebuilding companies are interested in accuracy of such housing prices prediction models as it directly impacts their activities relating to exposure management, hedging, arbitraging, investment and financing decisions. Policymakers frequently monitor housing prices to better understand their impact on local employment, revenue generation at county and state levels and potential for foreclosure and bail-outs. Nowadays, more attention is being focused on housing prices due to the increasing number of foreclosures and defaults on subprime mortgages.

The housing sector has been studied extensively in macroeconomics. Some studies such as the ones by Stockman and Tesar [1], Lane [2], Girouard & Bloundal [3] consider housing prices in a dynamic, general equilibrium model. Other studies have looked at housing prices in the context of "bubbles" though research has shown that "bubbles" are empirically difficult to validate. For instance, Driffill and Sola [4] do not find any distinction between bubbles and switching processes. In a study of property prices in Taiwan and stock movements, Chen [5] was unable to conclude the presence of a bubble. Ortalo-Magne and Rady [6,7,8,9] have done extensive studies on the interactions between housing prices, housing transactions, demographic change, income distribution changes and aggregate economic activity. Traditional models explaining house prices have used locational factors (Mok, Chan & Cho [10], Do, Wilbur & Short [11]), structural variables (Mok [12]; Do and Grudnitski [13]; Do, Wilbur and Short [11]) such as floor area, median gross income (Asabere and Huffman [14]).

Although many studies have been conducted on forecasting house prices and whether there is a housing 'bubble', no single study has exclusively dealt with short-run and long run forecasts developed through the use of econometric models. Consequently, there is a need to investigate the performance of these econometric models in forecasting house prices.

In this paper, we evaluate the efficiency of various house prices forecasting models. The objective of this study is to determine which models forecast efficiently, given the out-of-sample forecasted house prices and the actual median house prices in short (1 month), medium term (6 months) and long run (12 months) periods. The next section of this paper discusses the various models which are evaluated. This section is followed by a description of the data and the methodology used. Finally, the results are summarized in the concluding section.

House Prices Forecasting Models

Structural and time-series models have been typically used to forecast house prices. The only study that formally evaluated house price forecasts was by Clapp & Giaccotto [25] but their study did not focus on smoothing models. Their study also lacked focus in that it did not cover a long sample period and was restricted to sale transactions in Miami.

In this paper, we evaluate the performance of three major types of exponential smoothing models and a multivariate regression model: Simple exponential smoothing, Holt's two parameter exponential smoothing and Winter's three parameter exponential smoothing.

Exponential smoothing models are adaptable to adjustments to include trend and/or seasonality (Jarrett [26]). Further, they provide self-correcting forecasts with built-in adjustments that regulate forecast values by changing them after "learning" from direction of past errors. Limitations of exponential smoothing models include: these models are extremely sensitive to the specification of the smoothing constant and exponential smoothing techniques result in forecasts that lag behind the turning points in actual time series data. The first limitation can be overcome by using software which automatically computes the optimal smoothing constant, based on the objective of minimization of root mean square error. Forecasts generated by these optimal smoothing constants tend to be more accurate than those generated by regression models. The second limitation may not be so important, since turning points in any cycle tend to be more infrequent than not. Despite limitations of exponential smoothing models, a number of large-scale empirical studies, such as Reid [27], Newbold and Granger [28], Makridakis and Hibon [29], Makridakis et al. [30] and Mahmoud and Housseini [31] have used various exponential smoothing models, and have found them to perform quite well, *in the aggregate,* when compared with time series models.

Models, Data and Methodology

Simple Exponential Smoothing:

$$F_{t+1} = \alpha X_t + (1-\alpha)F_t \quad (1)$$

where F_{t+1} and F_t are the forecast values for the next and current periods, X_t is the current observed value, t is the current time period and α is the smoothing constant that takes that takes on values between 0 and 1.

Holt's two parameter exponential smoothing:

Three equations and two smoothing constants are used in this model:

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + b_{t-1}) \quad (2)$$

$$b_t = \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1} \quad (3)$$

$$F_{t+m} = S_t + b_t m \quad (4)$$

where b is the growth factor, S_t is the smoothed value in the current period, β is the new smoothing constant for the trend and m is the number of periods being forecasted.

Winter's three parameter exponential smoothing:

Four equations are necessary to forecast:

$$S_t = \alpha X_t/(I_{t-L}) + (1-\alpha)(S_{t-1} + b_{t-1}) \quad (5)$$

$$I_t = \beta X_t/S_t + (1-\beta)I_{t-L} \quad (6)$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \quad (7)$$
$$F_{t+m} = (S_t + b_t m)I_{t-L+m} \quad (8)$$

where S_t is the smoothed value in the current period, α is the smoothed constant for the data, X_t is the actual value in current period, I_t is the seasonality estimate, b_t is the trend estimate, γ is the smoothing constant for the trend estimate and L is the length of seasonality.

Multivariate Regression Model

The specific model that we use to estimate house prices forecast model is

$$\text{Ln (HP)}_t = \alpha + \beta\, u_{t-1} + \Delta\, i_{t-1} + €_t \quad (9)$$

where Ln $(HP)_t$ is the log of real median house prices in period t, u_{t-1} is the local unemployment rate in period t-1 and i_{t-1} is the real interest rate measured by the 30 year Federal Home loan bank series.

Data and Methodology:

We obtained monthly data for median house prices for Los Angeles County for the period January 1979 through January 2007 from the California Association of Realtors. We collected data on unemployment rates for the Los Angeles-Long Beach region from the Bureau of Labor Statistics and the interest rate data was obtained from the Federal Reserve Bank of Saint Louis.

In order to test whether our results will be different if we used state and national median home prices, we collected data on House Price Indices from the Office of Federal Housing Enterprise Oversight (http://www.ofheo.gov/HPI.asp). All models are estimated initially from January 1979 through January 2006. The estimated models are then used to make forecasts for the next 1 month (that is, for February 2006), 6 months (that is, for February 2006-July 2006), and 12 months (February 2006-January 2007). The 1-month, 6-month and 12-month forecasts are then compared with the actual median house prices or indices in the forecast window (February 2006 through January 2007).

Results

The simple exponential smoothing models outperformed all other forecasting models in all periods of forecasting. Forecasts generated by the multivariate regression models performed poorly at all levels. Our results indicate that forecasts of prices of houses are better made with fewer variables. Addition of more information through various variables in forecast models do not seem to significantly improve forecasts.

References, tables and appendices are available from the author upon request.