

# Boom–Bust Cycles and the Forecasting Performance of Linear and Non-Linear Models of House Prices

W. Miles

Published online: 11 July 2007  
© Springer Science + Business Media, LLC 2007

**Abstract** The tremendous rise in house prices over the last decade has been both a national and a global phenomenon. The growth of secondary mortgage holdings and the increased impact of house prices on consumption and other components of economic activity imply ever-greater importance for accurate forecasts of home price changes. Given the boom–bust nature of housing markets, nonlinear techniques seem intuitively very well suited to forecasting prices, and better, for volatile markets, than linear models which impose symmetry of adjustment in both rising and falling price periods. Accordingly, Crawford and Fratantoni (*Real Estate Economics* 31:223–243, 2003) apply a Markov-switching model to U.S. home prices, and compare the performance with autoregressive-moving average (ARMA) and generalized autoregressive conditional heteroscedastic (GARCH) models. While the switching model shows great promise with excellent in-sample fit, its out-of-sample forecasts are generally inferior to more standard forecasting techniques. Since these results were published, some researchers have discovered that the Markov-switching model is particularly ill-suited for forecasting. We thus consider other non-linear models besides the Markov switching, and after evaluating alternatives, employ the generalized autoregressive (GAR) model. We find the GAR does a better job at out-of-sample forecasting than ARMA and GARCH models in many cases, especially in those markets traditionally associated with high home-price volatility.

**Keywords** Bubbles · Forecasting models · Generalized autoregressive model · Nonlinear techniques

## Introduction

The boom in home values and the growing financial sophistication of consumers, lenders and pension funds have made modeling house prices more important than at

---

W. Miles (✉)

Department of Economics, Wichita State University, 1845 Fairmount, Wichita, KS 67260-0078, USA  
e-mail: [william.miles@wichita.edu](mailto:william.miles@wichita.edu)

any time in recent memory. Some suggest housing may, in some parts of the U.S., exhibit speculative bubbles (Case and Schiller 2003). Bubbles are accompanied by crashes, and thus lead to highly volatile price movements. The volatility of house prices has been found to be a significant determinant of the probability of default and mortgage prepayment (Foster and Van Order 1984, Crawford and Rosenblatt 1995). Moreover, while house prices have a large impact on the financial sector, they also have a greater effect on the real sector of the economy than at any recent time. An IMF study of housing booms in OECD countries from 1970–2000 found that in 19 of 20 cases, a crash in prices was followed by recession (Economist 2005). Consumption in the United States accounted for 90% of output growth in the first 3 years of the current recovery, and it has of course been closely tied to real estate prices. Indeed research from Goldman Sachs found that consumption in the United States was more closely linked to home prices than in other countries (Economist 2005).

It is thus vital to accurately forecast home prices. Since the housing market is subject to boom and bust cycles, nonlinear models, which allow for asymmetric adjustment according to whether prices have been rapidly rising or falling, offer the promise of better forecasts than standard linear techniques. Accordingly, Crawford and Fratantoni (2003) employ a Markov-switching (MS) model to house prices in five states, and compare forecasting results to linear autoregressive (AR)-moving average (MA) (ARMA) and generalized AR conditional heteroscedastic (GARCH) models. Initially, the MS models appear to have a lot of potential for predicting home prices, as the in-sample fit is generally much better than in the ARMA and GARCH cases. However, out-of-sample results reveals that the MS model is superior to the two alternatives in only three of 15 possible cases.

Since the Crawford and Fratantoni results were published, other researchers have discovered that the MS model provides particularly poor out-of-sample forecasts (Bessec and Bouabdallah 2005). At the same time the use of other nonlinear techniques besides the MS model has grown exponentially in many areas of finance. Thus we will employ an alternative nonlinear method, the generalized AR (GAR) model to the same states that Crawford and Fratantoni utilized. Our findings will indicate that the GAR is substantially better at forecasting home prices than the MS, especially in those markets known for high levels of volatility.

As noted, housing is more important for the economy than at any point in recent memory, so better housing forecasts will be useful. U.S. commercial banks have a greater proportion of real assets on their balance sheets than at any time in over 30 years. Better forecasts can help banks, as well as their regulators develop better capital requirements. Many institutional investors hold large amounts of mortgage-backed securities. Prepayment risk is a major concern, and Foster and Van Order find volatility to be a determinant of such risk. Better forecasts would help such institutions better manage such risk. Thus improved forecasts promise important benefits for many parties exposed to housing and mortgage activity.

## Previous Literature and Methodology

The wide swings in house prices in the U.S., especially in certain high-volatility states, has caused speculation that bubbles may be present in the housing market. A bubble is

generally defined by an asset price rising above its fundamental value. There is a large controversy in finance as to whether bubbles exist at all, for any asset. In an efficient market, bubbles should not exist, as arbitrageurs could presumably profit by selling overpriced assets. At the same time, many researchers have come to believe such inefficiencies do exist, and research on rational bubbles proceeds apace.

Flood and Hodrick (1990) review the literature on empirical tests for speculative bubbles in liquid markets such as those for equities and currency. Testing for bubbles in such a fashion involves specifying an asset pricing model, thus rejection of market efficiency could be due to a lack of efficiency or an incorrectly specified model of asset prices. The authors conclude that most apparent finding of bubbles may be due to this or other difficulties of interpreting empirical results.

Gatzlaff and Tirtiroglu (1995) review the literature on efficiency in the real estate market. They review evidence that the housing market may not be as efficient as more liquid financial assets such as equities. This is plausible, as housing entails greater information and transactions costs.

Case and Schiller (2003) use a unique survey to investigate the possibility of housing bubbles in three U.S. cities. The authors first find that housing fundamentals such as income, employment and interest rates explain much of the variation in prices in most states, but there are eight states in which the fundamentals explain less of house prices. In three cities – Boston, Los Angeles, and San Francisco – the authors believe there may well be bubbles, based on responses to survey questions.

Having defined a bubble in terms of expectations–homebuyers making purchases based on a belief in future price increases–the authors present results pointing to a bubble. In particular, people in boom cities often believe that prices can keep rising at implausibly high rates, and perceive little risk of price drops. The authors also note that the housing market is dominated by amateurs, rather than professional speculators. This could mean fewer arbitrageurs to damp bubbles.

Thus there may well be a propensity to bubbles in some parts of the country. Bubbles are of course followed by busts, in which owners are reluctant to sell, and time-on-the-market (TOM) increases. Clearly, house sellers face a tradeoff between setting a high asking price and suffering high TOM (Anglin et al. 2003). Empirically, however, research into TOM has been plagued by an inability to explain much of the variation in time to sale (Horowitz 1992; Knight 2002).

Ong and Koh (1990) find that expectations of capital gains lead to greater TOM, which is consistent with a glut of unsold houses in the aftermath of a burst bubble. It is also consistent with the pattern of asymmetric price adjustment, where house prices are sticky downward. As Case and Schiller (2003) ask, “What happens in a bust?...It is important to point out that the housing market is not an auction market. Prices do not fall to clear the market quickly, as one observes in most asset markets. Selling a home requires agreement between buyer and seller. It is a stylized fact about the housing market that bid–ask spreads widen when demand drops, and the number of transactions drops sharply. This must mean that sellers resist cutting prices.” (p. 335)

Thus prices should exhibit nonlinear patterns of adjustment, and forecasting models which take account of such nonlinearity may do better than those which impose symmetric adjustment to rising and falling prices. Forecasts of housing prices may be more important than at any time for the U.S. economy. As noted, consumption is tied more closely to home prices in the U.S. than in other countries.

Housing, both directly and indirectly, is responsible for two-fifths of the jobs created in the U.S. since 2001 (Economist 2005). Home builders would clearly like to be able to forecast prices to avoid unwanted build-ups in inventories. Banks would certainly like to forecast home prices for optimal portfolio management. Despite the growth of secondary mortgage markets, real estate represents 33.5% of the U.S. banking industries as of the 2006, the highest level since 1973 according to Federal Reserve officials (Wall Street Journal 2006).

There is an existing literature on home price forecasting using linear techniques. Case and Schiller (1989) was the canonical paper on forecasting house prices. The authors were investigating the informational efficiency of the housing market in the United States, and found that prices did not follow a random walk. Price changes are thus persistent and therefore forecastable. Gu (2002) confirms that different housing markets display varying autocorrelation patterns, and shows that in the case of California, excess returns are in principle possible using a simple trading rule. Gillen et al. (2001) find varying patterns of spatial autocorrelation in different housing markets in greater Philadelphia.

Zhou (1997) uses linear cointegration and error-correction models (ECM) to forecast prices for the U.S. Barot and Takala (2001) use cointegration analysis for house prices in Finland and Sweden.

Muellbauer and Murphy (1997) discuss boom and bust cycles in the UK housing market, and then go on to explicitly mention non-linearities in price changes, which they believe arise from transactions costs. However, the authors build a structural, rather than a forecasting model, the goal of which is to estimate elasticities and other parameters, rather than generate predictions about future price changes.

Thus the Crawford and Fratantoni (2003) paper is a major contribution. While the testing of hypotheses and estimation of structural parameters, as was the focus in Muellbauer and Murphy, is important, it is also obviously essential to generate accurate forecasts of future prices, given the role home prices now play for both the financial sector and the economy as a whole. And Crawford and Fratantoni is the first study to highlight the importance of the boom–bust nature of house prices, and to utilize a particular non-linear model in an attempt to capture this cycle and improve forecasts.

The authors employ quarterly Office of Federal Housing Enterprise and Oversight (OFHEO) house price index data from five states – California, Florida, Massachusetts, Ohio and Texas – over the period 1979–2001. They develop three different types of forecasting models for each state. The first two types are well-known-ARMA and GARCH models. While these types of time series models are very useful for producing forecasts, the authors note that housing prices in many markets exhibit boom–bust cycles. Thus prices may behave differently depending on what “state” housing conditions are in. Accordingly, the authors employ a MS model to capture this asymmetry.

The authors thus estimate ARMA, GARCH and MS models for all five states, with the number of parameters and lag orders determined by goodness of fit, or minimum root mean square error (RMSE), within the sample. Initially, results seem to be favorable for the nonlinear MS model, as it has far better fit than either alternative in all of the states. When out-of-sample forecasts are performed, however, results are very different. Three different forecast horizons – 2 years, 5 years and 10 years – are used for the five states, for a total of 15 different forecasts. Of the 15 forecasts, the MS model has the minimum RMSE in only three.

Thus the nonlinear MS model does not appear to improve forecast accuracy. However, since the Crawford–Fratantoni paper was published, other research has revealed that the MS model is particularly ill-suited for forecasting (Bessec and Bouabdallah 2005). The authors’ examination of the literature points to a very relevant finding for the Crawford–Fratantoni paper-MS models usually display superior in-sample fit but worst out-of-sample forecasting performance relative to linear models. Thus some nonlinear model may do a better job of capturing the boom–bust nature of the housing market than standard linear estimation, but the MS approach may not be the particular specification needed.

What nonlinear model, then should be used, if MS exhibits poor forecast performance? There has been tremendous growth in the last several years in the use of nonlinear models in economics and finance, as theory often suggests that a given series should exhibit asymmetric adjustment depending on past behavior, and hence nonlinearity. There are several prominent models, besides MS.

One of the most popular nonlinear models is the threshold autoregression (TAR). This allows the dependent variable,  $y_t$  to follow two (or more) different processes depending on whether the lagged value of  $y_t$  is above or below a threshold value. Formally, for the very simple case of an AR(1) model a TAR would be specified as:

$$y_t = a_1y_{t-1} + \varepsilon_{1t} \text{ if } y_{t-1} > \tau \tag{1}$$

$$y_t = a_2y_{t-1} + \varepsilon_{2t} \text{ if } y_{t-1} \leq \tau \tag{2}$$

Here  $\tau$  is the threshold, and  $\varepsilon$  is a random error. In trying to decide whether a linear AR(1) model, such as

$$y_t = ay_{t-1} + \varepsilon_t \tag{3}$$

is the correct specification, we can test the null hypothesis of  $a_1=a_2$ . If we cannot reject this null, we conclude that the process is linear and adjustment symmetric. However, standard F-tests cannot be employed, as the threshold  $\tau$  is chosen by a search procedure designed to minimize the sum of squared errors, and thus an F-test is biased in favor of finding TAR. Instead, a bootstrapping procedure developed by Hansen (1997) must be employed. We employed this test for the five state home price indices over the same 1979–2001 period as utilized in Crawford and Fratantoni, but were unable to reject the null of linearity for any state (results available upon request).

The TAR process is a very specific form of nonlinearity, so failure to find TAR effects does not suggest that a given process is linear. Another popular nonlinear technique is the bilinear (BL) model. In this model AR and MA terms are interacted, so that coefficients on past values are time-varying and depend on past shocks. As an example,

$$y_t = a_1y_{t-1} + b_1\varepsilon_{t-1} + c_1y_{t-1}*\varepsilon_{t-1} + \varepsilon_t \tag{4}$$

or, equivalently,

$$y_t = (a_1 + c_1\varepsilon_{t-1})*y_{t-1} + b_1\varepsilon_{t-1} + \varepsilon_t \tag{5}$$

So here the value of  $\varepsilon_{t-1}$  affects the impact of  $y_{t-1}$ . While intuitively, the BL model may be a compelling choice, Brunner and Hess (1995) find many undesirable

properties of this technique. In particular, standard optimization methods often miss the optimum estimates, and standard  $t$ -statistics often have non-standard distributions in finite samples.

Given these difficulties, the best nonlinear model may be the simplest, or most general—the GAR. This model starts, in a sense, as a linear AR model, and then adds squares, cubes and other higher powers, of the lagged dependent variable, as well as interactions between  $y_{t-i}$  and  $y_{t-j}$  for  $i \neq j$ . A typical GAR model might look as follows:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3}^2 + a_4 y_{t-1} * y_{t-2} + \varepsilon_t \quad (6)$$

Thus,  $\partial y_t / \partial y_{t-1} = a_1 + a_4 y_{t-2}$ . The effect of  $y_{t-1}$  thus depends on the past value of  $y$ . The model can simulate a large number of functional forms. And unlike the BL model, the GAR can be estimated with simple OLS. We will thus use the GAR to develop forecasting models for the same states that Crawford and Fratantoni employed, and compare the forecasting results to the best forecasting ARMA and GARCH models for each of the states.

## Data and Estimation

We will begin with the same OFHEO price indices for the same five states – California, Florida, Massachusetts, Ohio and Texas – that Crawford and Fratantoni employed, and for the same period, utilizing quarterly data from 1979 through 2001, for a total of 92 observations. We will later extend the analysis through the third quarter of 2005, yielding one-hundred-and-seven observations, and see if using more observations affects the results.

Crawford and Fratantoni find for each state one model – an ARMA, GARCH or MS – that produces optimal out-of-sample forecasts. As noted, the MS is best in only three of 15 cases. Through experimentation, we will develop GAR models for all five states, and pit these GAR models against the best ARMA and GARCH models that Crawford and Fratantoni obtained in a contest of out-of-sample forecasting performance. For the three cases where the MS model was best for Crawford and Fratantoni, we will compare the GAR's performance with both the best ARMA and GARCH specifications for the particular forecast horizon in the given state.

ARMA models have been a staple of forecasting for some years. The standard specification is as follows:

$$y_t = \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i} \quad (7)$$

Here, the AR portion of the model is represented by  $\sum_{i=1}^p \alpha_i y_{t-i}$  while the MA portion is denoted as  $\sum_{i=0}^q \beta_i \varepsilon_{t-i}$ , and the  $\varepsilon_{t-i}$  are lagged residuals. If one or more of the  $\alpha_i$  coefficients is equal to one, the process is nonstationary, and contains a unit root. Such a data generating process would be termed an AR *integrated* MA, or ARIMA process. However, testing by Crawford and Fratantoni has established that the OFHEO state price indices do not contain unit roots and are stationary.

The other set of forecasts, based on GARCH models, are similar to the ARMA specification in that they posit a constant rate of growth for the long run mean. However, GARCH processes exhibit time-varying volatility; that is a time-dependent variance. Some financial markets exhibit periods of relative tranquility, followed by periods of very large price swings. Miller and Peng (2006), find GARCH effects in about 17% of the MSA housing markets they survey in the U.S. The specification for a GARCH is:

$$y_t = \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \tag{8}$$

$$\varepsilon_t = v_t h_t^{0.5} \tag{9}$$

$$h_t = a_0 + \sum_{i=1}^q a_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta h_{t-i} \tag{10}$$

Here  $v_t$  is a random error with mean zero and a variance of 1, and  $h_t$  is the conditional variance process. While this standard GARCH specification is quite common, it requires non-negativity constraints on the parameters of Eq. 10, since a variance can never be negative. Nelson (1991) develops an alternative specification, called exponential GARCH (E-GARCH). The mean equation for an E-GARCH model is identical to Eq. 9, while a typical specification for the conditional variance is:

$$\ln(h_t) = a_0 + a_1 (\varepsilon_{t-1} / h_{t-1}^{0.5}) + b_1 |\varepsilon_{t-1} / h_{t-1}^{0.5}| + \beta \ln(h_{t-1}) \tag{11}$$

Since the conditional variance is in log-linear form, it is impossible for  $h_t$  to be negative, regardless of the value of  $\ln(h_t)$ . Thus some of the coefficients can be negative. Given that this model is less constrained than the standard GARCH, it has been found to out-forecast other GARCH models. Not surprisingly, the E-GARCH was found to perform well relative to other models for Crawford and Fratantoni.

We will measure forecasting performance by RMSE, as the authors did, and we will also compare performance in terms of mean absolute error (MAE). The MAE criterion is most appropriate when the cost of a forecast error rises proportionally with respect to the absolute size of the error. With RMSE, the cost of the error rises as the square of the error, and so large errors can be weighted far more than proportionally. This corresponds to a “quadratic loss function”, common in some theoretical models. Whether MAE or RMSE is most appropriate surely varies according to circumstances and individual institutions, and in any case we will find that the two measures pick the same model in all but several instances.

In addition, while RMSE and MAE are good relative measures, both depend on the scale of the forecast variable. Moreover, each could hypothetically be quite low, and still contain systematic bias and do a poor job of forecasting average price changes. Accordingly, we will examine how each model performs according to a measure known as the bias proportion (listed as bias in Tables 2 and 4). The bias proportion measures *systematic* error, or the extent to which the average values of the forecast and actual series differ from each other. Unlike RMSE and MAE, bias

proportion is scale-invariant. A given forecasting model may have a systematic positive or negative bias and do a poor job of tracking the actual mean of price changes, and measures such as RMSE and MAE could well miss this defect. We will thus evaluate forecasts based on the three performance measures of RMSE, MAE and bias.

Parameter estimates for the 1979–2001 period are displayed in Table 1, and results are displayed in Table 2. For the GAR specifications, a notation of GAR(1,3,1\*3) denotes a first and third AR lag, and the interaction of the first and third. The specifications of the ARMA and ARCH models are standard—i.e. ARMA(1,2) denotes one AR and two MA lags. Starting with the case of Ohio, Crawford and

**Table 1** Forecasting models

Coefficients		<i>P</i> values
<b>OHIO</b>		
AR(2)-GARCH(1,1)		
C	3.5	0.000
AR(1)	0.15	0.24
AR(2)	0.04	0.74
ARCH(1)	0.28	0.05
GARCH(1)	0.6	0.003
GAR (1,2,3, 1*1, 2*2, 3*2)		
C	1.3	0.109
AR(1)	-0.2	0.09
AR(2)	0.54	0.002
AR(3)	0.21	0.17
AR(1*1)	0.06	0.000
AR(2*2)	-0.0	0.003
AR(3*2)	-0.0	0.68
<b>TEXAS</b>		
AR(2) E-GARCH(1,1)		
C	3.6	0.000
AR(1)	0.24	0.001
AR(2)	0.15	0.03
GARCH(1)	-0.3	0.02
AbsARCH(1)	0.94	0.002
ARCH(1)	-0.3	0.103
ARMA(2,2)		
C	0.22	0.42
AR(1)	0.76	0.19
AR(2)	0.12	0.81
MA(1)	-0.4	0.41
MA(2)	-0.2	0.55
GAR(1,3, 1*1)		
C	1.3	0.05
AR(1)	0.4	0.000
AR(3)	0.21	0.02
AR(1*3)	-0.0	0.28
<b>MASS</b>		
AR(2) E-GARCH(1,1)		
C	11.9	0.003
AR(1)	0.63	0.000
AR(2)	0.25	0.01
GARCH(1)	1.0	0.000
AbsARCH(1)	-0.0	0.08
ARCH(1)	0.1	0.06



**Table 1** (continued)

Coefficients		<i>P</i> values
ARMA(2,2)		
C	0.8	0.343
AR(1)	0.2	0.29
AR(2)	0.67	0.001
MA(1)	0.42	0.04
MA(2)	−0.4	0.001
GAR(1,3, 1*3)		
C	0.74	0.33
AR(1)	0.61	0.000
AR(3)	0.33	0.002
AR(1*3)	−0.0	0.62
CAL		
AR(1)		
C	0.77	0.11
AR(1)	0.85	0.000
AR(1) ARCH(1)		
C	5.8	0.06
AR(1)	0.88	0.000
ARCH(1)	0.21	0.24
GAR (1, 1*1)		
C	0.77	0.12
AR(1)	0.85	0.000
AR(1*1)	0.00	0.99
FLA		
ARMA(1,1)		
C	2.5	0.001
AR(1)	0.38	0.01
MA(1)	−0.1	0.34
AR(1) E-GARCH(1,1)		
C	2.4	0.000
AR(1)	0.28	0.000
GARCH(1)	−0.7	0.000
AbsARCH(1)	0.4	0.002
ARCH(1)	0.13	0.27
GAR(1,3, 1*3)		
C	2.1	0.01
AR(1)	0.21	0.19
AR(3)	0.27	0.05
AR(1*3)	−0.0	0.99

Data 1979:1–2001:4

Fratantoni found that the best forecasting specification at all horizons was an AR(2)-E-GARCH(1,1) model. Although several GAR models were experimented with, the best fitting GAR(1,2,3, 1\*1, 2\*2, 3\*2) had worse out-of-sample performance for all horizons than the E-GARCH. Thus for Ohio, nonlinear modeling does not appear to improve predictive power.

For Texas, the AR(2)-E-GARCH(1,1) model found optimal by Crawford and Fratantoni was also best here at the 5-year horizon, as it outperforms the GAR(1,3,1\*1). However, at the 10-year horizon, the GAR(1,3,1\*1) outperforms the E-GARCH by the RMSE and MAE measures, while the E-GARCH still achieves minimum bias

**Table 2** Forecasting results

	RMSE	MAE	Bias
Ohio			
Out-of-sample forecast 1992:1–2001:4 (10 years)			
<b>AR(2) GARCH(1,1)</b>	<b>1.533</b>	<b>1.189</b>	<b>0.001896</b>
GAR(1,2,3, 1*1, 2*2, 3*2)	3.305	2.987	0.306376
Out-of-sample forecasts 1997:1–2001:4 (5 years)			
<b>AR(2) GARCH(1,1)</b>	<b>1.467</b>	<b>1.077</b>	<b>0.002979</b>
GAR(1,2,3, 1*1, 2*2, 3*2)	2.708	2.307	0.37329
Out-of-sample forecasts 2000:1–2001:4 (2 years)			
<b>AR(2) GARCH(1,1)</b>	<b>1.466</b>	<b>0.971</b>	<b>0.12366</b>
GAR(1,3, 1*1)	2.908	2.533	0.699501
Texas			
Out-of-sample forecasts 1992:1–2001:4 (10 years)			
AR(2) E-GARCH(1,1)	2.915	2.351	<b>0.006791</b>
<b>GAR(1,3, 1*1)</b>	<b>2.881</b>	<b>2.329</b>	0.018157
Out-of-sample forecasts 1997:1–2001:4 (5 years)			
<b>AR(2) E-GARCH(1,1)</b>	<b>2.751</b>	<b>2.042</b>	<b>0.40679</b>
GAR(1,3, 1*1)	2.958	2.251	0.52182
Out-of-sample forecasts 2000:1–2001:4 (2 years)			
ARMA(2, 2)	3.762	2.8	0.54648
<b>GAR(1,3, 1*1)</b>	<b>3.199</b>	<b>2.148</b>	<b>0.368069</b>
Mass			
Out-of-sample forecasts 1992:1–2001:4 (10 years)			
ARMA(2, 2)	4.744	3.632	0.039289
<b>GAR(1,3, 1*3)</b>	<b>4.244</b>	<b>3.337</b>	<b>0.000352</b>
Out-of-sample forecasts 1997:1–2001:4 (5 years)			
<b>AR(2) E-GARCH(1,1)</b>	<b>3.384</b>	<b>2.415</b>	<b>0.197193</b>
GAR(1,3, 1*1)	5.35	4.23	0.610318
Out-of-sample forecasts 2000:1–2001:4 (2 years)			
<b>AR(2) E-GARCH(1,1)</b>	<b>2.997</b>	<b>2.41</b>	<b>0.188074</b>
ARMA(2,2)	3.982	3.176	0.524215
GAR(1,3, 1*1)	3.274	2.471	0.289578
Cal			
Out-of-sample forecasts 1992:1–2001:4 (10 years)			
AR (1)	6.881	5.869	0.056165
<b>AR(1)-ARCH(1)</b>	<b>6.828</b>	<b>5.779</b>	0.073849
GAR(1,1, 1*1)	6.87	5.87	<b>0.054737</b>
Out-of-sample forecasts 1997:1–2001:4 (5 years)			
<b>AR(1)</b>	<b>6.032</b>	<b>4.982</b>	<b>0.66676</b>
GAR(1,1, 1*1)	6.047	4.998	0.667805
Out-of-sample forecasts 2000:1–2001:4 (2 years)			
<b>AR(1) ARCH(1)</b>	<b>6.037</b>	<b>5.196</b>	<b>0.662114</b>
GAR(1,1, 1*1)	6.31	5.45	0.703998
Fla			
Out-of-sample forecasts 1992:1–2001:4 (10 years)			
ARMA(1,1)	3.922	3.251	0.000221
AR(1) E-GARCH(1,1)	3.982	3.354	0.027199
<b>GAR(1,3, 1*3)</b>	<b>3.841</b>	<b>3.207</b>	<b>0.000019</b>
Out-of-sample forecasts 1997:1–2001:4 (5 years)			
AR(2) E-GARCH(1,1)	4.396	3.827	0.42712
<b>GAR(1,3, 1*3)</b>	<b>4.038</b>	<b>3.55</b>	<b>0.385216</b>
Out-of-sample forecasts 2000:1–2001:4 (2 years)			
ARMA(1,1)	5.217	4.9743	<b>0.909598</b>
<b>GAR(1,3, 1*3)</b>	<b>5.204</b>	<b>4.9752</b>	0.913932

Data 1979:1–2001:4. The RMSE and MAE numbers in bold are those with the lowest value among the competing forecasting models

proportion. At the 2-year horizon, the GAR has better forecasts than the ARMA(2,2) which Crawford and Fratantoni found worked best over that horizon.

In Massachusetts, the authors had found that the MS model worked best, though only at the 2-year horizon. Here, the AR(2) E-GARCH(1,1) forecasts better than GAR at 2 years, but at 10 years the GAR(1,3,1\*3) out-forecasts the ARMA(2,2) which was found by Crawford and Fratantoni to be optimal at that horizon. At 5 years, the E-GARCH is again optimal. Thus nonlinearity seems important for the Bay State in one of three horizons.

For California, the best model at 2 years is the ARCH(1) according to RMSE and MAE, but the GAR(1,1, 1\*1) achieves minimum bias proportion. At the 5-year horizon, the AR(1) model out-forecasts the GAR(1,1, 1\*1). At 10 years, the best forecasting model is the ARCH(1).

In Florida, the GAR model is the best predictor at all horizons, performing better in terms of RMSE and MAE than both the ARMA and GARCH models. The only exception is that at the 2-year horizon, the ARMA(2,2) model has a slightly lower bias proportion. Nonlinearity may therefore be a very important feature of home price adjustment in the sunshine state.

Given the results, how did the nonlinear GAR model perform as an alternative to the MS model? In Crawford–Fratantoni, the MS technique was the best forecasting model in three of 15 possible cases, using the RMSE criterion. Here, the GAR model had the lowest RMSE in six of 15 cases, while it also has the lowest MAE in six of 15. The GAR model displayed the lowest bias proportion in five of 15. This is a substantial improvement over the MS model.

Thus far, we have examined data only from 1979–2001. OFHEO data does not go back much further for many states, so this was as much as was available several years ago. However, we will add data through the third quarter of 2005, for a total of 107 observations. While more data are always desirable in any empirical study, it is especially the case with time series. The new parameter estimates are displayed in Table 3 and the results in Table 4.

For Ohio, there is still no horizon at which the GAR model yields the best forecast. In Texas, GAR yields the best forecasts at the 2-and-5-year horizons by RMSE and MAE, while the E-GARCH model is best at 10 years by all measures and at 5 years by the bias proportion. In Massachusetts, GAR produces the best predictions at 10 years by all metrics, while at 5 years, the MAE criterion would pick the E-GARCH, but RMSE and bias proportion would choose the GAR. At 2 years, the ARMA model is clearly superior to all others.

In Florida, GAR does not improve forecasts at the 2 or 10 year horizons. However, at the 5-year horizon, GAR yields a smaller MAE and bias proportion than forecasts generated by the E-GARCH model.

California is the housing market that most clearly benefits from nonlinear forecasts. The GAR model yields the best forecast at all horizons. California has experienced several pronounced boom and bust cycles over the past 25 years, and its market thus appears to benefit most from a nonlinear specification. Ohio, in contrast, has had a relatively stable housing market compared to coastal states, and thus the GAR model fails to improve forecast performance at any horizon. Altogether, with this longer time series, GAR provides the best forecasts in seven of 15 cases as measured by RMSE, as well as in seven of 15 cases by the MAE yardstick, and seven cases by bias proportion. It often takes a longer time series to detect the

**Table 3** Forecasting models

Coefficients		<i>P</i> values
<b>OHIO</b>		
AR(2)-GARCH(1,1)		
C	4.3	0.000
AR(1)	0.12	0.32
AR(2)	0.01	0.9
ARCH(1)	0.28	0.03
GARCH(1)	0.6	0.000
GAR (1,2,3, 1*1, 2*2, 3*2)		
C	1.4	0.05
AR(1)	-0.2	0.06
AR(2)	0.53	0.001
AR(3)	0.22	0.12
AR(1*1)	0.06	0.000
AR(2*2)	-0.0	0.001
AR(3*2)	-0.0	0.647
<b>TEXAS</b>		
AR(2) E-GARCH(1,1)		
C	3.7	0.000
AR(1)	0.32	0.001
AR(2)	0.057	0.61
GARCH(1)	1.01	0.000
AbsARCH(1)	-0.1	0.000
ARCH(1)	0.07	0.11
ARMA(2,2)		
C	0.25	0.35
AR(1)	0.76	0.17
AR(2)	0.12	0.79
MA(1)	-0.4	0.38
MA(2)	-0.2	0.51
GAR(1,3, 1*1)		
C	1.4	0.02
AR(1)	0.4	0.000
AR(3)	0.2	0.015
AR(1*3)	-0	0.23
<b>MASS</b>		
AR(2) E-GARCH(1,1)		
C	6.8	0.06
AR(1)	0.58	0.000
AR(2)	0.24	0.006
GARCH(1)	-0.6	0.02
AbsARCH(1)	-0.0	0.7
ARCH(1)	0.27	0.01
ARMA(2,2)		
C	0.79	0.3
AR(1)	0.24	0.16
AR(2)	0.64	0.000
MA(1)	0.35	0.04
MA(2)	-0.4	0.000
GAR(1,3, 1*3)		
C	0.83	0.27
AR(1)	0.54	0.000
AR(3)	0.37	0.000
AR(1*3)	-0.0	0.69

**Table 3** (continued)

Coefficients		<i>P</i> values
CAL		
AR(1)		
C	1.6	0.02
AR(1)	0.79	0.00
AR(1) ARCH(1)		
C	7	0.04
AR(1)	0.89	0.000
ARCH(1)	0.52	0.003
GAR (1, 1*1)		
C	1.0	0.12
AR(1)	1.12	0.000
AR(1*1)	-0.0	0.000
FLA		
ARMA(1,1)		
C	-0.1	0.43
AR(1)	1.04	0.000
MA(1)	-0.8	0.000
AR(1) E-GARCH(1,1)		
C	5.2	0.000
AR(1)	0.5	0.000
GARCH(1)	0.95	0.000
AbsARCH(1)	-0.2	0.000
ARCH(1)	0.12	0.016
GAR(1,3, 1*3)		
C	2.5	0.003
AR(1)	0.13	0.28
AR(3)	0.2	0.06
AR(1*3)	0.03	0.001

Data 1979:1–2005:3

inherent nonlinearity in a variable, and thus to benefit from nonlinear forecasts. Nonlinear modeling appears even more important for the data through 2005, especially in states known for sometimes volatile housing markets.

### Conclusions

Nonlinear models have been widely applied in most areas of finance in the last several years to obtain better forecasts, and account for inherently nonlinear adjustment processes. These techniques have vastly increased understanding of the workings of foreign exchange, bond and equity markets. Housing markets, however, have rarely been modeled with such techniques, with Crawford and Fratantoni being a very important exception. Yet given the relatively illiquid nature of housing, as compared to financial securities, and the corresponding boom–bust nature of housing markets, home prices are one area that can benefit substantially from nonlinear forecasting techniques.

**Table 4** Forecasting results

	RMSE	MAE	Bias
Ohio			
Out-of-sample forecasts 1995:4–2005:3 (10 years)			
<b>AR(2) GARCH(1,1)</b>	<b>1.353</b>	<b>1.036</b>	<b>0.01463</b>
GAR(1,2,3, 1*1, 2*2, 3*2)	3.131	2.724	0.159
Out-of-sample forecasts 2000:4–2005:3 (5 years)			
<b>AR(2) GARCH(1,1)</b>	<b>1.439</b>	<b>1.107</b>	<b>0.009743</b>
GAR(1,2,3, 1*1, 2*2, 3*2)	2.605	2.085	0.168
Out-of-sample forecasts 2003:4–2005:3 (2 years)			
<b>AR(2) GARCH(1,1)</b>	<b>1.365</b>	<b>1.046</b>	<b>0.0152</b>
GAR(1,3, 1*1)	2.787	2.041	0.44
Texas			
Out-of-sample forecasts 1995:4–2005:3 (10 years)			
<b>AR(2) E-GARCH(1,1)</b>	<b>2.418</b>	<b>1.897</b>	<b>0.036</b>
GAR(1,3, 1*1)	2.483	1.913	0.072
Out-of-sample forecasts 2000:4–2005:3 (5 years)			
AR(2) E-GARCH(1,1)	2.416	1.708	<b>0.0409</b>
<b>GAR(1,3, 1*1)</b>	<b>2.352</b>	<b>1.667</b>	0.046
Out-of-sample forecasts 2003:4–2005:3 (2 years)			
ARMA(2, 2)	2.525	2.036	0.559
<b>GAR(1,3, 1*1)</b>	<b>2.073</b>	<b>1.701</b>	<b>0.482</b>
Mass			
Out-of-sample forecasts 1995:4–2005:3 (10 years)			
ARMA(2, 2)	5.343	3.875	0.296
<b>GAR(1,3, 1*3)</b>	<b>4.819</b>	<b>3.572</b>	<b>0.1639</b>
Out-of-sample forecasts 2000:4–2005:3 (5 years)			
AR(2) E-GARCH(1,1)	4.57	<b>2.999</b>	0.12
<b>GAR(1,3, 1*1)</b>	<b>4.27</b>	3.242	<b>0.05</b>
Out-of-sample forecasts 2003:4–2005:3 (2 years)			
AR(2) E-GARCH(1,1)	6.974	4.661	0.351
<b>ARMA(2,2)</b>	<b>6.41</b>	<b>4.29</b>	<b>0.258</b>
GAR(1,3, 1*1)	6.602	4.46	0.279
Cal			
Out-of-sample forecasts 1995:4–2005:3 (10 years)			
AR (1)	9.504	6.331	0.1652
AR(1)-ARCH(1)	9.401	6.719	0.247
<b>GAR(1,1, 1*1)</b>	<b>8.173</b>	<b>6.029</b>	<b>0.0381</b>
Out-of-sample forecasts 2000:4–2005:3 (5 years)			
<b>AR(1)</b>	<b>12.139</b>	8.145	0.377
GAR(1,1, 1*1)	<b>9.346</b>	<b>5.918</b>	<b>0.0709</b>
Out-of-sample forecasts 2003:4–2005:3 (2 years)			
AR(1) ARCH(1)	16.076	12.791	0.633
<b>GAR(1,1, 1*1)</b>	<b>14.038</b>	<b>10.182</b>	<b>0.506</b>
Fla			
Out-of-sample forecasts 1995:4–2005:3 (10 years)			
ARMA(1,1)	10.153	7.386	0.421
<b>AR(1) E-GARCH(1,1)</b>	<b>9.004</b>	<b>6.228</b>	<b>0.228</b>
GAR(1,3, 1*3)	9.206	6.332	0.268
Out-of-sample forecasts 2000:4–2005:3 (5 years)			
<b>AR(1) E-GARCH(1,1)</b>	<b>12.207</b>	9.61	0.619
GAR(1,3, 1*3)	12.28	<b>9.485</b>	<b>0.595</b>
Out-of-sample forecasts 2003:4–2005:3 (2 years)			
<b>ARMA(1,1)</b>	<b>11.57</b>	<b>10.31</b>	<b>0.793</b>
GAR(1,3, 1*3)	17.235	16.002	0.862

Data 1979:1–2005:3. The RMSE and MAE numbers in bold are those with the lowest value among the competing forecasting models

Results here indicate that, for the data extending into 2005, the GAR model is very important, especially in those states which exhibit the greatest house price volatility. These results suggest that models such as GAR may not add much forecastability in relatively stable markets such as Ohio (Ohio had the lowest raw variability, as measured by the ratio of the standard deviation to the mean of prices, of any of the states in the sample). On the other hand, nonlinear methods appear invaluable for markets such as California, with a history of housing bubbles and crashes. As rapidly rising and volatile house prices are a phenomenon in a growing number of localities and regions around the globe, nonlinear modeling will be ever more helpful in predicting the direction of future prices.

**Acknowledgments** I thank the Barton School of Business at Wichita State University for research support, and Stan Longhofer of the Center for Real Estate at Wichita State University, without whom this paper would not have been possible.

## References

- Anglin, P., Rutherford, R., & Springer, T. (2003). The trade-off between the selling price of residential properties and time-on-the-market: The impact of price setting. *Journal of Real Estate Finance and Economics*, 26, 95–111.
- Barot, B., & Takala, K. (2001). Prices and inflation: A cointegration analysis for Finland and Sweden. In K. Takala (Ed.), *Studies in time series analysis of consumption and asset prices*. Helsinki: Bank of Finland.
- Bessec, M., & Bouabdallah, O. (2005). What causes the forecasting failure of Markov-switching models? A Monte Carlo study. *Studies in Nonlinear Dynamics and Econometrics*, 9.
- Brunner, A., & Hess, G. (1995). Potential problems in estimating bilinear time series models. *Journal of Economic Dynamics and Control*, 19, 663–681.
- Case, K., & Schiller, R. (1989). The efficiency of the market for single-family homes. *American Economic Review*, 79, 125–137.
- Case, K., & Shiller, R. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity*, 2, 299–362.
- Crawford, G., & Fratantoni, M. (2003). Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. *Real Estate Economics*, 31, 223–243.
- Crawford, G., & Rosenblatt, E. (1995). Efficient mortgage default option exercise: Evidence from loan loss severity. *The Journal of Real Estate Research*, 10, 543–555.
- Economist (2005). In come the waves. *Economist*, 66–68, June 18, 2005.
- Flood, R., & Hodrick, R. (1990). On testing for speculative bubbles. *Journal of Economic Perspectives*, 4, 85–101.
- Foster, C., & Van Order, R. (1984). An option-based model of mortgage default. *Housing Finance Review*, 3, 351–372.
- Gatzlaff, D., & Tirtiroglu, D. (1995). Real estate market efficiency: issues and evidence. *Journal of Real Estate Literature*, 3, 157–189.
- Gillen, K., Thibodeau, T., & Wachter, S. (2001). Anisotropic autocorrelation in house prices. *Journal of Real Estate Finance and Economics*, 23, 5–30.
- Gu, A. (2002). The predictability of home prices. *Journal of Real Estate Research*, 24, 213–234.
- Hansen, B. (1997). Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics*, 2, 1–14.
- Horowitz, J. (1992). The role of the list price in housing markets: Theory and an econometric model. *Journal of Applied Econometrics*, 7, 115–129.
- Knight, J. (2002). Listing price, time on market, and ultimate selling price: causes and effects of listing price changes. *Real Estate Economics*, 30, 213–237.
- Miller, N., & Peng, L. (2006). Exploring metropolitan housing price volatility. *Journal of Real Estate Finance and Economics*, 33, 5–18.

- 
- Muellbauer, J., & Murphy, A. (1997). Booms and busts in the U.K. *Housing Market Economic Journal*, 107, 1701–1727.
- Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59, 347–369.
- Ong, S., & Koh, Y. (1990). Time-on-market and price trade-offs in high-rise housing sub-markets. *Urban Studies*, 37, 2057–2071.
- Wall Street Journal (2006). Housing begins to pinch nation's banks. *Wall Street Journal*, September 1, 2006, p C1.
- Zhou, Z. (1997). Forecasting sales and price for existing single-family homes: a VAR model with error correction. *Journal of Real Estate Research*, 14, 155–167.