



# Forecasting Real Housing Price Growth in the Eighth District States

David E. Rapach and Jack K. Strauss

The authors consider forecasting real housing price growth for the individual states of the Federal Reserve's Eighth District. They first analyze the forecasting ability of a large number of potential predictors of state real housing price growth using an autoregressive distributed lag (ARDL) model framework. A number of variables, including the state housing price-to-income ratio, state unemployment rate, and national inflation rate, appear to provide information that is useful for forecasting real housing price growth in many Eighth District states. Given that it is typically difficult to determine a priori the particular variable or small set of variables that are the most relevant for forecasting real housing price growth for a given state and time period, the authors also consider various methods for combining the individual ARDL model forecasts. They find that combination forecasts are quite helpful in generating accurate forecasts of real housing price growth in the individual Eighth District states. (JEL C22, C53, E37)

Federal Reserve Bank of St. Louis *Regional Economic Development*, 2007, 3(2), pp. 33-42.

**T**he rollercoaster ride of the housing market continues to receive considerable attention in the popular and financial press. There is currently speculation of a precipitous drop in housing prices in certain regions of the country after the sharp rise in housing prices (“bubble”?) over the past decade. Policymakers are keenly interested in housing price fluctuations and their potential impact on household consumption spending, as evinced by numerous comments by former Federal Reserve Chairman Alan Greenspan and current Chairman Ben Bernanke. This interest appears warranted: The median household now holds more of its wealth in housing than in stocks and has greater access to cash through refinancing backed by housing wealth (Greenspan and Kennedy, 2005).

Given the substantial interest in housing price fluctuations, the present paper investigates fore-

casts of real housing price growth in the individual states of the Federal Reserve's Eighth District (Arkansas, Illinois, Indiana, Kentucky, Missouri, Mississippi, and Tennessee). We focus on forecast horizons of four and eight quarters because these horizons are relevant to forecasting over the business cycle, and most recent discussions of housing price fluctuations focus on possible swings in housing prices over business-cycle horizons.<sup>1</sup> We consider a large number of potential predictors (25) of real housing price growth for each state. This is motivated by a sizable literature that examines the determinants of housing prices using in-sample

---

<sup>1</sup> The literature on forecasting housing prices in the United States at the aggregate or state level is relatively sparse, especially compared with the massive literature on forecasting economic variables such as U.S. gross domestic product (GDP) and inflation. The extant literature on forecasting housing prices in the United States tends to focus on long-run trends (Hendershott and Weicher, 2002).

---

David E. Rapach is an associate professor of economics and Jack K. Strauss is the Simon Professor of Economics at Saint Louis University. The authors acknowledge financial support from the Simon Center for Regional Forecasting at Saint Louis University. The authors thank participants at the 2007 Eighth District Business and Economics Research Group (BERG) conference and a referee for very helpful comments. The results reported in this paper were generated using GAUSS 6.0; the GAUSS programs are available at <http://pages.slu.edu/faculty/rapachde/Research.htm>.

© 2007, The Federal Reserve Bank of St. Louis. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

tests; see, for example, Cho (1996), Abraham and Hendershott (1996), and Johnes and Hyclak (1999). Potential determinants of housing prices in this literature include income, interest rates, construction costs, and labor market variables such as the unemployment rate and size of the labor force.<sup>2</sup>

Following Stock and Watson (1999, 2003, and 2004), we generate simulated out-of-sample forecasts of real housing price growth using an autoregressive distributed lag (ARDL) model framework. More specifically, when forecasting real housing price growth for a given Eighth District state, we estimate 25 individual ARDL forecasting models of real housing price growth, where each ARDL model includes one of the potential predictors. This provides a convenient framework for analyzing the forecasting ability of each of the individual potential predictors of real housing price growth.

The plethora of potential predictors of real housing price growth also leads us to consider combination forecasts. Typically, it is difficult to identify a priori the particular variable (or small set of variables) that is most relevant for forecasting a variable such as real housing price growth, especially because the predictive ability of individual variables can vary markedly over time.<sup>3</sup> Combination forecasts provide a way of incorporating information that may be useful for forecasting in environments with a large number of potential predictors, and they have been shown to work well in a number of recent forecasting applications involving GDP growth, inflation, and employment growth; see, for example, Stock and Watson (1999, 2003, and 2004) and Rapach and Strauss (2005 and 2007). We consider a number of different methods for combining the individual ARDL model forecasts from the extant literature and investigate their ability to help generate reliable forecasts of

real housing price growth in the Eighth District states.<sup>4</sup>

Previewing our results, we find that a number of the individual predictors are able to improve on forecasts of real housing price growth relative to an autoregressive (AR) benchmark model, sometimes substantially. These variables include the housing price-to-income ratio, state unemployment rate, and national inflation rate. However, there is no single variable that is able to improve on the AR model forecasts across all of the Eighth District states at all of the forecast horizons considered, and there are instances where a variable that performs very well for one particular state performs poorly for another. Fortunately, we also find that some of the forecast combining methods perform quite well and almost always provide sizable improvements in forecast accuracy relative to the AR benchmark model.

The rest of the paper is organized as follows: The next section outlines the econometric methodology, and the third section presents the empirical results.

## ECONOMETRIC METHODOLOGY

Let  $\Delta y_t = y_t - y_{t-1}$ , where  $y_t$  is the log level of real housing prices at time  $t$ . Furthermore, let

$$y_{t+h}^h = (1/h) \sum_{j=1}^h \Delta y_{t+j},$$

so that  $y_{t+h}^h$  is the (approximate) growth rate of real housing prices from time  $t$  to  $t+h$ ;  $h$  is the forecast horizon. Let  $x_{i,t}$  ( $i = 1, \dots, n$ ) represent one of  $n$  potential predictors of real housing price growth. An individual ARDL model based on the predictor  $x_{i,t}$  is given by

$$(1) \quad y_{t+h}^h = \alpha + \sum_{j=0}^{q_1-1} \beta_j \Delta y_{t-j} + \sum_{j=0}^{q_2-1} \gamma_j x_{i,t-j} + \varepsilon_{t+h}^h,$$

where  $\varepsilon_{t+h}^h$  is an error term. Equation (1) can be used to construct a set of recursive (expanding estimation window) simulated out-of-sample forecasts of  $y_{t+h}^h$  using information available at time  $t$ ,

<sup>2</sup> We focus on real housing price growth at the state level in this paper primarily because it allows us to examine regional differences in housing price fluctuations while still having a fairly large number of potential predictors available at the state level. Although state-level housing prices are able to capture some important geographic differences in housing price fluctuations, as we mention in the conclusion, we are also planning to investigate forecasts of real housing price growth for individual metropolitan areas in the Eighth District in future research.

<sup>3</sup> See Stock and Watson (2003) for evidence of this in the context of forecasting U.S. GDP growth and inflation.

<sup>4</sup> See Timmermann (2006) for a recent survey of forecast combining methods.

and we denote the forecast of  $y_{t+h}^h$  formed at time  $t$  for a given predictor  $x_{i,t}$  by  $\hat{y}_{i,t+h|t}^h$ . More specifically,  $\hat{y}_{i,t+h|t}^h$  is calculated by plugging  $\Delta y_{t-j}$  ( $j = 0, \dots, q_1 - 1$ ) and  $x_{i,t-j}$  ( $j = 0, \dots, q_2 - 1$ ) into (1), with the  $\alpha$ ,  $\beta_j$ , and  $\gamma_j$  parameters set equal to their ordinary least squares (OLS) estimates based on data available from the start of the sample through period  $t$  and  $\varepsilon_{t+h}^h$  set to its expected value of zero. We select the lag lengths ( $q_1$  and  $q_2$ ) in (1) using the Schwarz information criterion (SIC) and a minimum value of zero for  $q_1$  and one for  $q_2$  (to ensure that the potential predictor  $x_{i,t}$  appears in (2)) and a maximum value of four for  $q_1$  and  $q_2$ .<sup>5</sup> Dividing the total sample into in-sample and out-of-sample portions of size  $R$  and  $P$ , respectively, we use this procedure to generate a series of  $P - (h - 1)$  recursive simulated out-of-sample forecasts for the ARDL model that includes  $x_{i,t}$ , which we denote as

$$\left\{ \hat{y}_{i,t+h|t}^h \right\}_{t=R}^{T-h} .^6$$

Note that the lag lengths  $q_1$  and  $q_2$  are selected anew when forming each out-of-sample forecast, so that the lag lengths for the ARDL forecasting model are allowed to vary through time. In our applications in the next section, we consider 25 potential predictors, and so we will have 25 series of  $h$ -step-ahead individual ARDL model forecasts of real housing price growth for each of the seven states in the Eighth District.

We also compute recursive simulated out-of-sample forecasts for an AR model, which is given by (1) with the restriction  $\gamma_j = 0$  ( $j = 0, \dots, q_2 - 1$ ) imposed. The series of out-of-sample forecasts are generated using a procedure analogous to that for the ARDL forecasting model described above.<sup>7</sup> Following much of the forecasting literature, the

AR model serves as a benchmark forecasting model.

We consider three types of methods for combining the individual ARDL model forecasts. Some of the combining methods require a holdout period to calculate the weights ( $\{w_{i,t}\}_{i=1}^n$ ) used to combine the individual ARDL model forecasts, and we use the first  $P_0$  observations from the out-of-sample period as the initial holdout period. This leaves us with a total of  $P - (h - 1) - P_0$  out-of-sample forecasts available for evaluation.<sup>8</sup> In our applications in the next section, we evaluate the benchmark AR model, individual ARDL model, and combination forecasts over the 1995:Q1–2006:Q4 out-of-sample period. Importantly, this period includes the bull housing market that has prevailed in many parts of the country over the past decade.

The first type of combining method uses simple schemes: mean, median, and trimmed mean. The mean (median) combination forecast is simply the average (median) of the individual ARDL model forecasts, while the trimmed mean combination forecast takes the average of the individual ARDL model forecasts after dropping the highest and lowest individual ARDL model forecasts. Stock and Watson (1999 and 2003) find that simple combinations of individual ARDL model forecasts consistently outperform an AR benchmark model (although by a fairly limited margin) with respect to forecasting U.S. real GDP growth and inflation.<sup>9</sup>

The second type of combining procedure we employ uses a discount mean square forecast error (DMSFE) criterion over the holdout out-of-sample period to determine the weights used to combine the individual ARDL model forecasts formed at time  $t$ ; see Stock and Watson (2004). More specifically, the DMSFE combining method uses the weights

<sup>5</sup> The SIC and the Akaike information criterion (AIC) are two popular model selection procedures. Note that we obtain similar results when we select the lag lengths in (1) using the AIC.

<sup>6</sup> Note that the first forecast uses all data available at time  $R$  to form a forecast of  $y_{R+h}^h$ ; this forecast is denoted by  $\hat{y}_{i,R+h|R}^h$ . The information set is then updated by one period, and we use all data available at time  $R + 1$  to form a forecast of  $y_{(R+1)+h}^h$ ; this second forecast is denoted by  $\hat{y}_{i,(R+1)+h|R+1}^h$ . We continue in this manner through the end of the out-of-sample period, leaving us with  $P - (h - 1)$  recursive simulated out-of-sample forecasts,  $\{\hat{y}_{i,t+h|t}^h\}_{t=R}^{T-h}$ .

<sup>7</sup> We select the lag length ( $q_1$ ) for the AR model using the SIC and a minimum (maximum) value of zero (four) for  $q_1$ .

<sup>8</sup> Note that we use the first  $P_0$  observations from the out-of-sample period to estimate the combining weights used to generate the first combination forecast available for evaluation. We then use the first  $P_0 + 1$  observations from the out-of-sample period to estimate the combining weights used to generate the second combination forecast available for evaluation. We continue in this manner through the end of the available out-of-sample period, leaving us with a series of  $P - (h - 1) - P_0$  out-of-sample combination forecasts available for evaluation.

<sup>9</sup> The simple combining methods obviously do not require a holdout period, as the combining weights are not estimated.

$$w_{i,t} = m_{i,t}^{-1} / \sum_{j=1}^n m_{j,t}^{-1} \quad (i = 1, \dots, n),$$

where

$$m_{i,t} = \sum_{s=R}^{t-h} \theta^{t-h-s} (y_{s+h}^h - \hat{y}_{i,s+h|s}^h)^2$$

and the parameter  $\theta$  is a discount factor. When  $\theta = 1$ , there is no discounting, whereas  $\theta < 1$  means that greater importance is attached to the recent forecasting performance of the individual ARDL models in determining the combining weights. In the next section, we consider  $\theta$  values of 1.0 and 0.9 in our applications.

The final type of combining method we use is the “cluster” approach recently developed by Aiolfi and Timmermann (2006) based on their  $C(K, PB)$  algorithm. The initial cluster combination forecast is generated by first grouping the individual ARDL model forecasts over the holdout out-of-sample period,

$$\left\{ \hat{y}_{i,s+h|s}^h \right\}_{s=R}^{R+(P_0-1)-(h-1)} \quad (i = 1, \dots, n),$$

into  $K$  equal-sized clusters based on the MSFE, with the first cluster containing the individual ARDL model forecasts with the lowest MSFE values, the second cluster containing the individual ARDL model forecasts with the next lowest MSFE values, and so on. The initial combination forecast is the average of the individual ARDL model forecasts contained in the first cluster. To form the second combination forecast, the MSFE is computed for the individual ARDL model forecasts,

$$\left\{ \hat{y}_{i,s+h|s}^h \right\}_{s=R+1}^{R+(P_0-1)-(h-1)+1} \quad (i = 1, \dots, n),$$

and the individual ARDL model forecasts are again grouped into  $K$  clusters based on the MSFE. The second combination forecast is again the average of the individual forecasts in the first cluster. We can proceed in this manner through the end of the available out-of-sample period to construct the complete set of combination forecasts. Following Aiolfi and Timmermann (2006), we consider  $K$  values of two and three in our applications in the next section.

## EMPIRICAL RESULTS

### Data

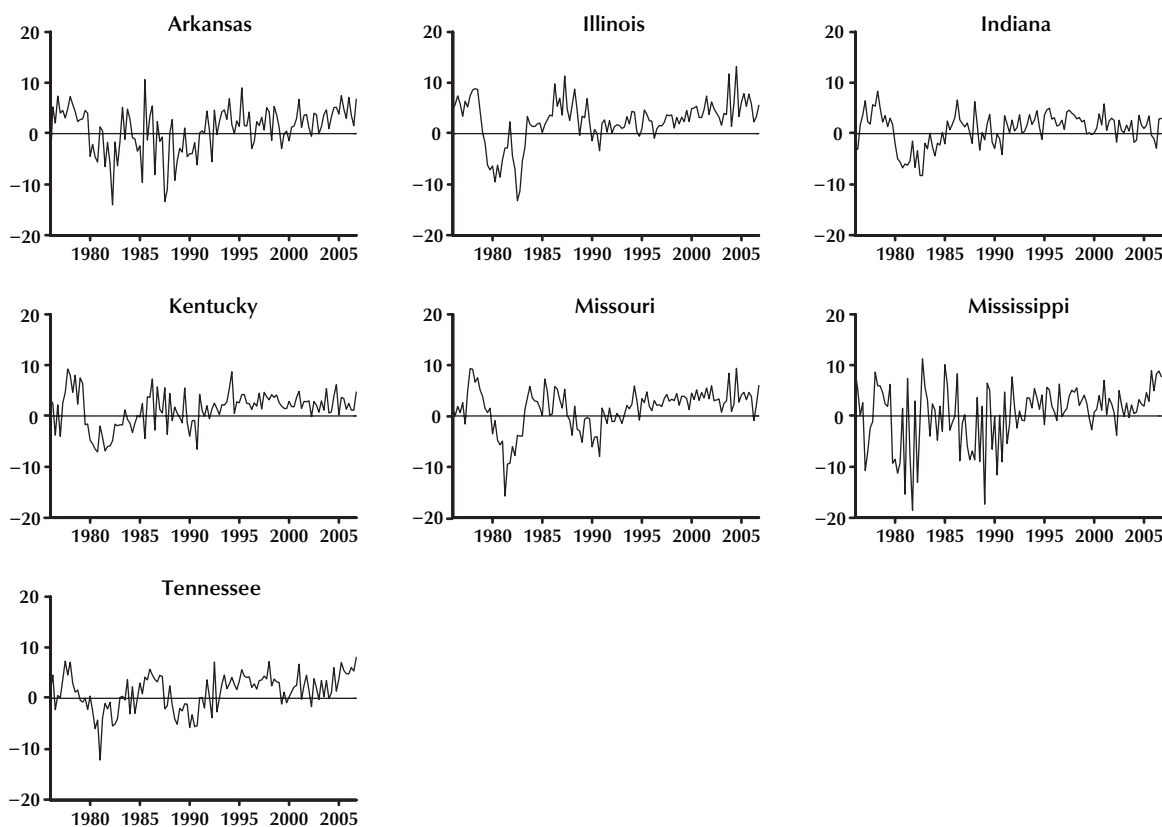
Nominal housing price indices for individual U.S. states starting in 1975:Q1 are available from Freddie Mac. The Conventional Mortgage Home Price Index provides a means for measuring the typical price inflation for houses within the United States using matched transactions on the same property over time to account for quality changes. Freddie Mac uses data from both purchase and refinance-appraisal transactions, and its database consists of over 33 million homes. The available sample for the housing price indices ends in 2006:Q4. We convert the nominal housing price index into real terms using the personal consumption expenditure (PCE) deflator from the Bureau of Economic Analysis (BEA). We then compute annualized growth rates as 400 times the differences in the log levels of real housing prices. The annualized real housing price growth rates are plotted in Figure 1. Note that real housing price growth is predominantly positive over much of our 1995:Q1–2006:Q4 out-of-sample forecast evaluation period, indicating that the individual states of the Eighth District typically experienced fairly strong housing markets over the past decade.<sup>10</sup>

As discussed above, we consider 25 potential predictors of real housing price growth for each state. Six of these are state-level variables:

- Ratio of housing price to per capita personal income
- Real per capita personal income
- Population
- Employment
- Labor force
- Unemployment rate

Nominal personal income data are from the BEA and are converted into per capita terms using

<sup>10</sup> The housing price indices exhibit exaggerated saw-tooth patterns in the first part of the sample for a number of the states. This appears to be an artifact of the development and construction of the housing price indices. To minimize the influence of these patterns when estimating the forecasting models, we smooth the real housing price growth observations up to 1984:Q4 by taking a moving average of the current and three previous real housing price observations. Smoothing of the early observations has been applied to the real housing price growth rate series depicted in Figure 1.

**Figure 1****Annualized Real Housing Price Growth, 1976:Q1–2006:Q4**

population data from the U.S. Census Bureau and then into real terms using the PCE deflator. The labor market variables are from the Bureau of Labor Statistics (BLS). The housing price-to-income ratio is a popular “valuation ratio” for housing prices that may help to signal whether housing is over- or under-valued. The income and employment variables provide measures of the ability of households to purchase housing and are thus potentially important determinants of housing demand. Significant changes in population can also lead to sizable shifts in housing demand.

We also consider five regional variables as predictors:

- Housing starts
- Building permits

- Homes for sale
- Homes sold
- Housing vacancy rate

These variables, all from the U.S. Census Bureau, are available for each of the four U.S. Census regions.<sup>11</sup> These housing market variables may provide signals of trends and supply conditions in housing markets that affect housing prices.

Finally, 14 national variables also serve as predictors:

- Average weekly hours in manufacturing

<sup>11</sup> Reflecting their U.S. Census Bureau classification, we use variables from the South region for Arkansas, Kentucky, Mississippi, and Tennessee and variables from the Midwest region for Illinois, Indiana, and Missouri. Note that these variables are not available at the state level for the entire sample period we consider.



## Rapach and Strauss

- Average weekly initial claims for unemployment insurance
- Manufacturers' new orders for consumer goods and materials (in chained 1982 dollars)
- Vendor performance
- Manufacturers' new orders of nondefense capital goods (in chained 1982 dollars)
- S&P 500 stock price index
- Real M2 money supply (in chained 2000 dollars)
- Term spread (10-year Treasury bond yield minus the federal funds rate)
- Consumer confidence index
- PCE deflator
- Industrial production
- Commercial and industrial loans outstanding (in chained 2000 dollars)
- Consumer installment credit outstanding
- Real effective mortgage rate

The first nine national predictors comprise nine of the ten leading economic indicators from the Conference Board<sup>12</sup>: These indicators potentially measure broad economic trends that can affect the demand for housing. Data on industrial production, commercial and industrial loans outstanding, and consumer installment credit outstanding are all from the Conference Board. These variables include credit measures that also may significantly influence housing prices. The nominal effective mortgage rate is from Freddie Mac, and we subtract the inflation rate based on the PCE deflator to approximate a real effective mortgage rate. The mortgage rate is an important component of the “user cost” of housing and thus a potentially important determinant of housing demand.

All of the predictors are transformed in an effort to render them stationary. This involves taking the first differences of log levels, with the following exceptions: We use levels for the unemployment rate, housing vacancy rate, unemployment claims, vendor performance, term spread, and consumer confidence; we use log levels for

the housing price-to-income ratio; and we use first differences for average weekly hours.

### **AR Benchmark and Individual ARDL Model Forecasting Results**

Table 1 reports forecasting results for the AR benchmark and individual ARDL forecasting models for each state. The table reports the MSFE for the AR benchmark model and the ratio of the individual ARDL model MSFE to the AR benchmark model MSFE. A ratio below unity thus indicates that the individual ARDL model has a lower MSFE than the AR benchmark. Results are reported for forecast horizons of four ( $h = 4$ ) and eight ( $h = 8$ ) quarters.

An important result in Table 1 is that no single predictor has an MSFE ratio that is below unity in all states for both forecast horizons; that is, there is no single predictor that delivers consistently more accurate forecasts than the benchmark AR model across all of the Eighth District states and both forecast horizons. The PCE deflator (inflation rate) produces an MSFE ratio below unity for all seven states at both horizons, with one exception (Indiana at  $h = 4$ ), and many of the MSFE ratios for the inflation rate are well below unity (for example, 0.26 for Kentucky at  $h = 8$ ), indicating substantial reductions in forecast accuracy relative to the AR model. The state housing price-to-income ratio—as mentioned above, a popular valuation ratio for housing—also performs quite well for Arkansas, Indiana, Kentucky, Mississippi, and Tennessee, with MSFE ratios all below unity (often substantially so) at both forecast horizons. However, the MSFE ratios for the state housing price-to-income ratio are well above unity for both horizons for Illinois and above unity for Missouri at the eight-quarter horizon. Other predictors that perform well for a number of Eighth District states are the state unemployment rate and consumer confidence, but there again are situations where the MSFE ratios for these variables are considerably above unity.

Looking at the results in Table 1 on a state-by-state basis, the state housing price-to-income ratio and state unemployment rate stand out for Arkansas. These predictors generate reductions in MSFE relative to the AR benchmark model of up to 31 percent and 54 percent at the four- and eight-

<sup>12</sup> The leading indicator we omit is national building permits, as this is included as a regional predictor.

**Table 1**  
**Individual ARDL Model Forecast Results: 1995:Q1–2006:Q4 Out-of-Sample Forecast Evaluation Period**

Predictor	Arkansas		Illinois		Indiana		Kentucky		Missouri		Mississippi		Tennessee	
	h = 4	h = 8	h = 4	h = 8	h = 4	h = 8	h = 4	h = 8	h = 4	h = 8	h = 4	h = 8	h = 4	h = 8
AR MSFE	4.90	5.49	4.66	5.66	2.15	1.93	1.18	2.25	2.93	4.69	9.70	9.44	3.43	4.56
State housing price-to-income ratio	0.70	0.46	1.89	2.74	0.84	0.69	0.67	0.38	0.92	1.13	0.58	0.35	0.76	0.52
State real personal income	1.01	1.02	1.02	1.03	0.95	1.13	0.99	1.01	1.14	1.43	1.02	1.01	1.05	1.06
State population	1.82	1.45	0.88	0.84	0.81	0.50	0.78	0.70	0.90	0.96	1.02	0.75	1.02	1.04
State employment	1.41	1.36	1.07	1.09	1.17	1.31	1.98	1.47	1.38	1.60	1.23	1.18	1.33	1.16
State labor force	1.14	1.08	1.05	0.71	1.13	1.28	1.07	0.98	1.17	1.17	1.08	1.01	1.29	1.16
State unemployment rate	0.69	0.49	1.39	1.89	0.69	0.47	0.67	0.69	1.21	2.05	0.99	0.89	1.15	1.34
Regional housing starts	1.08	0.97	1.18	1.06	1.29	1.13	1.15	1.01	1.24	1.22	0.99	0.94	1.11	0.95
Regional building permits	1.06	0.99	1.07	1.02	1.11	1.01	1.10	0.99	1.12	1.14	1.03	0.95	1.04	1.01
Regional homes for sale	1.36	1.21	1.00	0.97	1.29	1.09	1.33	1.17	1.05	1.05	1.18	1.01	1.22	1.08
Regional homes sold	1.02	1.02	1.00	1.00	1.06	1.00	1.97	0.65	1.00	1.00	1.15	1.02	1.08	0.86
Regional vacancy rate	1.12	1.18	0.82	0.57	1.17	1.23	0.85	0.62	1.11	0.83	1.08	1.10	0.95	0.83
Average weekly hours	1.33	1.10	1.58	1.23	1.20	1.18	1.40	1.20	1.08	1.16	1.16	1.08	1.17	1.11
Unemployment claims	0.95	1.02	1.03	1.02	0.90	0.67	0.97	0.91	0.97	1.04	0.98	0.96	0.96	0.97
New orders–consumer goods	1.14	1.04	1.31	1.13	1.06	1.10	1.17	1.29	1.03	1.12	1.14	1.05	1.08	1.06
Vendor performance	0.90	0.93	0.93	1.01	0.96	0.94	1.16	1.07	1.03	1.09	1.11	0.91	1.05	0.97
New orders–capital goods	1.00	1.00	1.01	1.00	1.13	1.02	1.01	1.01	1.01	1.02	1.03	1.00	1.01	1.00
S&P 500 index	1.05	1.04	1.08	1.10	1.01	1.02	1.05	1.03	1.03	1.03	1.06	1.04	1.01	1.02
Real M2	0.94	0.95	0.98	0.75	1.12	1.04	0.95	0.89	0.77	0.80	0.92	0.90	1.02	0.94
Term spread	0.81	0.82	1.05	1.10	1.25	1.76	2.00	1.96	1.16	1.65	1.30	1.01	1.23	1.29
Consumer confidence	0.87	0.82	1.09	1.16	1.03	1.20	0.75	0.54	0.67	0.44	0.88	0.56	0.94	0.72
PCE deflator (inflation rate)	0.81	0.64	0.83	0.58	1.08	0.96	0.68	0.26	0.53	0.32	0.63	0.43	0.89	0.62
Industrial production	1.05	1.04	1.21	1.11	0.99	0.93	1.05	1.08	1.07	1.13	1.03	0.98	1.08	1.00
Commercial/industry loans	1.28	0.98	1.07	0.92	1.21	1.26	1.08	1.02	1.02	0.99	1.04	1.01	1.16	1.02
Consumer credit outstanding	1.23	1.08	1.06	1.01	0.93	0.99	1.17	1.12	1.33	1.38	1.37	1.22	1.30	1.12
Real effective mortgage rate	0.95	1.01	1.15	1.37	1.03	1.07	1.03	1.05	0.97	1.29	1.11	1.13	1.12	1.19

NOTE: Entries in the AR MSFE row report the MSFE for the AR benchmark model. Entries in the other rows report the ratio of the MSFE for the individual ARDL model that includes the predictor indicated in the first column to the MSFE for the AR benchmark model.

**Table 2**  
**Combination Forecast Results: 1995:Q1–2006:Q4 Out-of-Sample Forecast Evaluation Period**

Combining method	Arkansas		Illinois		Indiana		Kentucky		Missouri		Mississippi		Tennessee	
	<i>h</i> = 4	<i>h</i> = 8	<i>h</i> = 4	<i>h</i> = 8	<i>h</i> = 4	<i>h</i> = 8	<i>h</i> = 4	<i>h</i> = 8	<i>h</i> = 4	<i>h</i> = 8	<i>h</i> = 4	<i>h</i> = 8	<i>h</i> = 4	<i>h</i> = 8
Mean	0.96	0.91	0.93	0.82	0.87	0.75	0.86	0.76	0.89	0.95	0.97	0.88	0.98	0.92
Median	0.99	0.99	0.97	0.97	0.93	0.89	0.90	0.91	0.97	1.03	1.00	0.97	1.00	0.99
Trimmed mean	1.03	0.98	1.02	0.94	0.89	0.78	0.98	0.84	0.97	1.02	1.01	0.89	1.18	0.99
DMSFE, $\theta = 1.0$	0.92	0.88	0.93	0.83	0.86	0.69	0.85	0.71	0.90	0.88	0.93	0.84	0.99	0.95
DMSFE, $\theta = 0.9$	0.90	0.82	0.91	0.76	0.88	0.68	0.84	0.69	0.86	0.76	0.89	0.80	0.99	0.93
<i>C</i> (2, <i>PB</i> )	0.90	0.85	0.92	0.81	0.88	0.69	0.81	0.61	0.86	0.84	0.92	0.81	0.98	0.86
<i>C</i> (3, <i>PB</i> )	0.87	0.76	0.85	0.72	0.84	0.63	0.76	0.52	0.80	0.70	0.86	0.74	1.00	0.79

NOTE: Entries report the ratio of the MSFE for the combining method indicated in the first column to the MSFE for the AR benchmark model.



quarter horizons, respectively. For Illinois, the state population, regional vacancy rate, and inflation rate display the best performance, with reductions in MSFE relative to the AR benchmark of up to 18 percent and 43 percent at the reported horizons. The state housing price-to-income ratio, state population, state unemployment rate, and unemployment claims produce large reductions in MSFE for Indiana, with reductions up to 31 percent and 53 percent at the two horizons. For Kentucky, six predictors are able to generate sizable reductions in MSFE relative to the AR benchmark at both horizons: the state housing price-to-income ratio, state population, state unemployment rate, regional vacancy rate, inflation rate, and consumer confidence. The largest reductions in MSFE are 33 percent and 74 percent at the four- and eight-quarter horizons, respectively. Real M2, consumer confidence, and the inflation rate lead to sizable reductions in MSFE relative to the AR benchmark for Missouri, with reductions of up to 47 percent and 68 percent at the two horizons. Three variables stand out for Mississippi: the state housing price-to-income ratio, consumer confidence, and inflation rate. The state housing price-to-income ratio leads to the largest reductions in MSFE (42 percent and 65 percent) at the two reported horizons. For Tennessee, the state housing price-to-income ratio and inflation rate lead to the largest reductions in MSFE relative to the AR benchmark at both of the reported horizons (up to 24 percent and 48 percent).

### **Combining Method Forecasting Results**

Table 2 reports the combination forecast results in the form of the ratio of the combining method MSFE to the AR benchmark MSFE, so that (as in Table 1) a ratio below unity indicates that the combining method forecast is more accurate than the AR benchmark forecast in terms of MSFE. The results in Table 2 show that the simple combining methods often produce reductions in MSFE relative to the AR benchmark of around 10 percent, and this is in line with the findings of Stock and Watson (1999 and 2003) in the context of U.S. GDP growth and inflation forecasts. The DMSFE combining method forecasts appear to perform somewhat better than the simple combining method forecasts in most cases, with the DMSFE combining method

based on  $\theta = 0.9$  leading to reductions in MSFE relative to the AR benchmark of approximately 10 to 15 percent at the four-quarter horizon and approximately 20 to 30 percent at the eight-quarter horizon in most cases. The cluster combining methods exhibit the best overall performance, especially the  $C(3, PB)$  method. With one exception (Tennessee at  $h = 4$ ), the MSFE ratios are all well below unity for the  $C(3, PB)$  method, with reductions in MSFE of up to 24 percent and 48 percent relative to the benchmark AR model at horizons of four and eight quarters, respectively (both for Kentucky). The  $C(3, PB)$  cluster combining method leads to average reductions in MSFE relative to the AR benchmark model across the seven states of approximately 15 percent and 30 percent at horizons of four and eight quarters, respectively. Given that it will be difficult to identify a priori the particular predictors that are most relevant for a given out-of-sample period, the performance of the combining methods—especially the  $C(3, PB)$  method—indicates that they provide a useful way of producing relatively accurate forecasts of real housing price growth in the Eighth District states in the presence of many potentially relevant predictors.

## **CONCLUSION**

We examine the ability of a host of economic variables to forecast real housing price growth for the seven individual states in the Federal Reserve's Eighth District. A number of variables, such as the state housing price-to-income ratio, state unemployment rate, consumer confidence, and inflation rate, produce forecasts that often substantially outperform a benchmark AR model in terms of MSFE in individual Eighth District states, but no single variable is able to improve on the AR benchmark for all states at all reported horizons. Given that it will be difficult to identify a priori the particular variable or small set of variables that are best suited for forecasting real housing price growth for a given state and time period, we also analyze the performance of forecast combining methods. We find that combining methods generally offer useful means of incorporating and culling information from a large number of potential predictors

## Rapach and Strauss

when forecasting real housing price growth in the Eighth District states.

Finally, we briefly discuss two ways that we are extending the research presented in this paper. First, we are currently applying the approaches employed in the present paper to a greater number of individual U.S. states, including larger U.S. states (in terms of population) that have experienced substantial increases in real housing prices over the past decade—states for which there are serious concerns of a housing price “bubble.” We are also preparing to apply the approach used in the present paper to forecasting real housing price growth for individual metropolitan areas in the Eighth District, as households are often interested in forecasts of housing price growth in their more immediate vicinity.

## REFERENCES

- Aiolfi, Marco and Timmermann, Allan. “Persistence in Forecasting Performance and Conditional Combination Strategies.” *Journal of Econometrics*, November-December 2006, 135(1-2), pp. 31-53.
- Abraham, Jesse M. and Hendershott, Patric H. “Bubbles in Metropolitan Housing Markets.” *Journal of Housing Research*, 1996, 7(2), pp. 191-207.
- Cho, Man. “House Price Dynamics: A Survey of Theoretical and Empirical Issues.” *Journal of Housing Research*, 1996, 7(2), pp. 145-72.
- Greenspan, Alan and Kennedy, James. “Estimates of Home Mortgage Originations, Repayments, and Debt on One-to-Four Family Residences.” Finance and Economic Discussion Series Paper 2005-41, Federal Reserve Board, September 2005.
- Hendershott, Patric H. and Weicher, John C. “Forecasting Housing Markets: Lessons Learned.” *Real Estate Economics*, Spring 2002, 30(1), pp. 1-11.
- Johnes, Geraint and Hyclak, Thomas. “House Prices and Regional Labor Markets.” *Annals of Regional Science*, February 1999, 33(1), pp. 33-49.
- Rapach, David E. and Strauss, Jack K. “Forecasting Employment Growth in Missouri with Many Potentially Relevant Predictors: An Analysis of Forecast Combining Methods.” Federal Reserve Bank of St. Louis *Regional Economic Development*, 2005, 1(1), pp. 97-112.
- Rapach, David E. and Strauss, Jack K. “Forecasting U.S. Employment Growth Using Forecast Combining Methods.” *Journal of Forecasting*, 2007 (forthcoming).
- Stock, James H. and Watson, Mark W. “Forecasting Inflation.” *Journal of Monetary Economics*, October 1999, 44(2), pp. 293-335.
- Stock, James H. and Watson, Mark W. “Forecasting Output Growth and Inflation: The Role of Asset Prices.” *Journal of Economic Literature*, September 2003, 41(3), pp. 788-829.
- Stock, James H. and Watson, Mark W. “Combination Forecasts of Output Growth in a Seven-Country Data Set.” *Journal of Forecasting*, September 2004, 23(6), pp. 405-30.
- Timmermann, Allan. “Forecast Combinations,” in Graham Elliott, Clive W.J. Granger, and Allan Timmermann, eds., *Handbook of Economic Forecasting*. Amsterdam: Elsevier, 2006, pp. 135-96.