Could We Have Predicted the Recent Downturn in Home Sales of the Four US Census Regions?♠

Rangan Gupta* , Christian K. Tipoy ♣ and Sonali Das#

Abstract

This paper analyzes the ability of a random walk and, classical and Bayesian versions of autoregressive, vector autoregressive and vector error correction models in forecasting home sales for the four US census regions (Northeast, Midwest, South, West), using quarterly data over the period of 2001:Q1 to 2004:Q3, based on an in-sample of 1976:Q1 till 2000:Q4. In addition, we also use our models to predict the downturn in the home sales of the four census regions over the period of 2004:Q4 to 2009:Q2, given that the home sales in all the four census regions peaked in 2005:Q3. Based on our analysis, we draw the following conclusions: (i) Barring the South, there always exists a Bayesian model which tends to outperform all other models in forecasting home sales over the out-of-sample horizon; (ii) When we expose our classical and 'optimal' Bayesian forecast models to predicting the peaks and declines in home sales, we find that barring the South again, our models did reasonably well in predicting the turning point exactly at 2005:Q3 or with a lead. In general, the fact that different models produce the best forecasting performance for different regions, highlights the fact that economic conditions prevailing at the start of the out-of-sample horizon are not necessarily the same across the regions, and, hence, vindicates our decision to look at regions rather than the economy as a whole. In addition, we also point out that there is no guarantee that the best performing model over the out-of-sample horizon is also well-suited in predicting the downturn in home sales.

1. INTRODUCTION

That there is a strong link between the housing market and economic activity in the US has been reported in a large number of recent papers (Green 1997, Iacoviello 2005, Case *et al.* 2005, Leamer 2007, Iacoviello and Neri 2010, Vargas-Silva 2008a,b, Ghent 2009, Ghent and Owyang 2009, amongst others). Leamer (2007) goes as far as suggesting that housing *is* the business cycle, with housing affecting the economy both macroeconomic and microeconomic activities. On one hand, with housing representing a large share of the total economy, movements in the housing sector spill over to the entire macro economy through new construction, renovations of existing property and the volume of home sales. Moreover, as indicated recently by Gupta *et al.* (forthcoming), housing responds significantly to adjustments in the interest rate, which peaks (and reaches a trough) according to business cycle peaks (and troughs). On the other hand, at the microeconomic level,

performance of financial institutions and real estate firms depend crucially on housing market activity, as the recent financial crisis would vouch for. In other words, the housing market affects the business cycle and is, in turn, affected by it. Moreover, as indicated by Dua and Smyth (1995), Dua and Miller (1996) and Dua *et al.* (1999), the housing market responds more quickly and more strongly to macroeconomic fluctuations relative to the average sector of an economy, since housing is a durable and flexible good, besides the fact that home construction does not require large firms. As such, timely and accurate forecasts of home sales can be invaluable to policy makers and financial institutions and real estate professionals.

Against this backdrop, our paper looks at the ability of a random walk (RW), and both the classical and the Bayesian versions of autoregressive (AR), vector autoregressive (VAR) and vector error correction (VEC) models in forecasting home sales for the four US census regions (Northeast, Midwest, South, West). For this prupose, we use quarterly data over the period of 2001:Q1 to 2004:Q3, based on an in-sample of period of 1976:Q1 to 2000:Q4. Note that the choice of the in-sample period, especially, the starting date, depends on data availability. The starting point of the out-of-sample period follows Iacoviello and Neri (2010), who observed a spurt of growth in the U.S. housing market, both in terms of home sales and prices, at the beginning of the 21st century, and a decline thereafter. We choose the end-point of the horizon as 2004:Q3, as we also use our models to predict the downturn in the home sales in the four census regions (over 2004:Q4 till 2009:Q2) and thus, stop (a year) prior to the date where the turning point actually occurred. In our case, the home sales in all the four census regions peaked in 2005:Q3, as depicted in Figure 1.

[INSERT FIGURE 1]

The motivation of our work is based on mainly three studies, namely those of Dua and Smyth (1995), Dua and Miller (1996) and Dua *et al.* (1999). While, Dua and Miller (1996) used Bayesian VAR (BVAR) models to forecast home sales for the state of Connecticut, Dua and Smyth (1995) and Dua *et al.* (1999) used the same to predict home sales for the aggregate US economy. In their model, Dua and Smyth (1995) included home sales, price of homes, mortgage interest rate, real disposable income and unemployment rate. Dua and Miller (1996) extended this model by including a leading index for the Connecticut economy and showed that the modified version was capable of producing improved forecast relative to the benchmark model of Dua and Smyth (1995). Finally, Dua *et al.* (1999) capitalized on this observation for the state of Connecticut and extended the model described in Dua and Smyth (1995) by adding six different leading indicators, namely, housing permits authorized, housing starts, the US Department of Commerce's composite index of eleven leading indicators, the short- and long-leading indices developed by the Center for International Business Cycle Research (CIBCR) at Columbia University and the leading index constructed by CIBCR that focussed solely on employment related variables. Dua *et al.* (1999) observed that the benchmark BVAR model, which included home sales, price of homes, mortgage rate, real personal disposable income, unemployment rate, supplemented by the building permits authorized as the leading indicator consistently produced the most accurate forecasts. Thus, following Dua *et al.* (1999), our multivariate forecasting models for each of the four census regions, comprise of home sales, price of homes, mortgage rate, real personal disposable income, unemployment rate and building permits authorized.[1] Understandably, the univariate models only include home sales, as it is the variable of interest. Note, the decision to look at the four regions of the US economy, rather than the aggregate US economy, as in Dua *et al.* (1999), emanates from the fact that economic conditions prevailing at a specific point of time, say for instance at the start of the out-of-sample

horizon, are not necessarily the same across the regions[2] (Carlino and DeFina 1998, 1999, Vargas-Silva 2008b, Gupta *et al.* (forthcoming)), and hence, as we show below, there cannot exist a single model that would tend to forecast best for all the four regions of the economy.

Our study thus extends the work of Dua *et al.* (1999) by not only looking at the four census regions, but by also investigating the forecasting ability of more detailed versions of the classical and Bayesian VAR models, namely the corresponding VEC models. Note, if the variables included in the VAR are non-stationary and share a common trend, not accounting explicitly for this comovement renders the VAR models as misspecified. The VECM, allowing for both short-and long-run dynamics, can thus be considered as an econometrically richer version, at least theoretically, relative to the VAR. In addition, this study is also the first in its attempt to analyze the ability of these models to predict the recent downturn of home sales in the four census regions of the US. However, moving to regional level analysis has its drawbacks as well: First, we have to use quarterly rather than monthly data as used by Dua *et al.* (1999), since information on real disposable personal income at the regional level, obtained by aggregating data for the states that fall under the respective census regions, is only available at quarterly frequencies. Second, information on the national level leading indexes used by Dua *et al.* (1999) is limited to only housing permits authorized and housing starts at the regional levels.[3] The rest of the paper is organized as follows: Section 2 lays out the basics of the classical and Bayesian variants of the VAR and VEC models, while subsections in Section 3 presents a discussion of the data, evaluation of alternative forecasting models and their ability to predict the recent downturn in the home sales of the four census regions. Finally, Section 4 concludes.

2. VAR, VEC, BVAR, and BVEC: Specification and Estimation[4]

Generally, economy-wide forecasting models are in the form of simultaneous-equations structural models. However, two problems often encountered with such models are as follows: (*i*) the correct number of variables need to be excluded for proper identification of individual equations in the system, which are however often based on little theoretical justification (Cooley and LeRoy, 1985), and; (*ii*) given that projected future values are required for the exogenous variables in the system, structural models are poorly suited for forecasting.

The vector autoregressive (VAR) model, though 'atheoretical', is particularly useful for forecasting exercises. Moreover, as shown by Zellner (1979) and Zellner and Palm (1974), any structural linear model can be expressed as a VAR moving average (VARMA) model, with the coefficients of the VARMA model being combinations of the structural coefficients. Under certain conditions, a VARMA model can be expressed as a VAR and a VMA model. Thus, a VAR model can be visualized as an approximation of the reduced-form simultaneous equation structural model.

Following Sims (1980), we can write an unrestricted VAR model as follows:
$$y_t = A_0 + A(L) y_t + \varepsilon_t,^5 \tag{1}$$
where y equals a $(n \times 1)$ vector of variables, which, in our case, includes home sales, price of homes, mortgage rate, real personal disposable income, unemployment rate and building permits authorized; $A(L)$ equals a $(n \times n)$ polynomial matrix in the backshift operator L with lag length p, and ε equals an $(n \times 1)$ vector of error terms. In our case, we assume that $\varepsilon \sim N(0, \sigma^2 I_n)$, where I_n equals a $(n \times n)$ identity matrix.

With cointegrated (non-stationary) series, we can transform the standard VAR model into a VEC model. The VEC model builds into the specification the cointegration relations, so that they restrict the long-run behavior of the endogenous variables to converge to their long-run, cointegrating relationships, while at the same time describing the short-run dynamic adjustment of the system. The cointegration terms, known as the error correction terms, gradually correct through a series of partial short-run adjustments.

Focusing on the practical case of y_t being a vector of n time series that are integrated[6] of order one, (i.e., $I(1)$),[7] then the error-correction counterpart of the VAR model in equation (1) converts into a VEC model as follows.[8]

$$\Delta y_t = \pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-1} + \varepsilon_t \qquad (2)$$

where $\pi = -[I - \sum_{i=1}^{p} A_i]$ and $\Gamma_i = -\sum_{j=i+1}^{p} A_j$.

Thus, a VECM is a restricted VAR designed for use with non-stationary series that are known to be cointegrated. While allowing for short-run adjustment dynamics, the VECM has cointegration relations built into the specification so that it restricts the long-run behaviour of the endogenous variables to converge to their cointegrating relationships. The cointegration term is known as the error correction term because the deviation from long-run equilibrium is corrected through a series of partial short-run adjustments, gradually.

VAR and VEC models typically use equal lag lengths for all variables in the model, which implies that the researcher must estimate many parameters, including many that prove statistically insignificant. This over-parameterization problem can create multicollinearity and a loss of degrees of freedom, leading to inefficient estimates, and possibly large out-of-sample forecasting errors. Some researchers exclude lags with statistically insignificant coefficients. Alternatively, researchers use near VAR models, which specify unequal lag lengths for the variables and equations.

Litterman (1981), Doan *et al.*, (1984), Todd (1984), Litterman (1986), and Spencer (1993), use a Bayesian VAR (BVAR) model to overcome the over-parameterization problem. Rather than eliminating lags, the Bayesian method imposes restrictions on the coefficients across different lag lengths, assuming that the coefficients of longer lags may approach more closely to zero than the coefficients on shorter lags. If, however, stronger effects come from longer lags, the data can override this initial restriction. Researchers impose the constraints by specifying normal prior distributions with zero means and small standard deviations for most coefficients, where the standard deviation decreases as the lag length increases. The first own-lag coefficient in each equation is the exception, which has a unitary mean. Finally, Litterman (1981) imposes a diffuse prior for the constant. We employ this 'Minnesota prior' in our analysis, where we implement Bayesian variants of the classical VAR and VEC models.
Formally, the means and variances of the Minnesota prior take the following form:

$$\beta_i \sim N(1, \sigma_{\beta_i}^2) \text{ and } \beta_j \sim N(0, \sigma_{\beta_j}^2) , \qquad (3)$$

where β_i equals the coefficients associated with the lagged dependent variables in each equation of the VAR model (i.e., the first own-lag coefficient), while β_j equals any other coefficient. In sum, the

4

prior specification reduces to a random-walk with drift model for each variable, if we set all variances to zero. The prior variances, $\sigma^2_{\beta_i}$ and $\sigma^2_{\beta_j}$, specify uncertainty about the prior means $\bar{\beta}_i = 1$, and $\bar{\beta}_j = 0$, respectively.

Doan *et al.,* (1984) propose a formula to generate standard deviations that depend on a small numbers of hyper-parameters: *w, d*, and a weighting matrix *f(i, j)* to reduce the over-parameterization in the VAR and VEC models. This approach specifies individual prior variances for a large number of coefficients, using only a few hyper-parameters. The specification of the standard deviation of the distribution of the prior imposed on variable *j* in equation *i* at lag *m*, for all *i, j* and *m*, equals S*(i, j, m)*, defined as follows:

$$S(i,j,m) = [w \times g(m) \times f(i,j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j}, \tag{4}$$

where *f(i, j) = 1*, if *i = j* and k_{ij} otherwise, with $(0 \le k_{ij} \le 1)$, and *g(m)* = m^{-d}, with *d > 0*. The estimated standard error of the univariate autoregression for variable *i* equals $\hat{\sigma}_i$. The ratio $\hat{\sigma}_i / \hat{\sigma}_j$ scales the variables to account for differences in the units of measurement, and hence, causes specification of the prior without consideration of the magnitudes of the variables. The term *w* indicates the overall tightness and equals the standard deviation on the first own lag, with the prior getting tighter as the value falls. The parameter *g(m)* measures the tightness on lag *m* with respect to lag 1, and equals a harmonic shape with decay factor *d*, which tightens the prior at longer lags. The parameter *f(i, j)* equals the tightness of variable *j* in equation *i* relative to variable *i*, and by increasing the interaction (i.e., the value of k_{ij}), we loosen the prior.[9]

The Bayesian variants of the classical VARs and VECMs are estimated using Theil's (1971) mixed estimation technique, which involves supplementing the data with prior information on the distribution of the coefficients. In an artificial way, the number of observations and degrees of freedom are increased by one, for each restriction imposed on the parameter estimates. The loss of degrees of freedom due to over-parameterization associated with a VAR model is, therefore, not a concern in the Bayesian models.

3. Models of Forecasting Home Sales for the Four US Census Regions:

3.1. Data and Prior Parameterization:

We first estimate the alternative univariate and multivariate models under both classical and Bayesian assumptions for the in-sample period of 1976:Q1 to 2000:Q4. We then compute the out-of-sample one- through four-quarters-ahead forecasts for the period of 2001:Q1 to 2004:Q3. As discussed before, the variables included in the multivariate model, for each of the four regions, are home sales, price of homes, mortgage rate, real personal disposable income, unemployment rate and building permits authorized. All data are obtained in their seasonally adjusted forms[10] in order to, *inter alia*, address the fact that the Minnesota prior is not well suited for seasonal data (Hamilton, 1994:362). The sources of the data used are as follows: Home sales are measured by the volume of existing single-family home sales, while home prices are measured by the median sales price of existing single-family homes and both of them come from the National Association of Realtors. The

national level mortgage interest rate is measured by the contract interest rate on single-family existing home purchases and is provided by the Federal Housing Finance Agency. Real personal disposable income is measured in billions of chained 2005 dollars and aggregated across the different states that falls under a specific census region. The unemployment rate of a specific region is the average of the unemployment rates of all the states that belong to a census region, and is equal to the civilian rate of sixteen years and over. These two series comes from the US Bureau of Economic Analysis. Leading indicators signal the start and end of a recession and are thus capable of predicting the path of future economic activity. In our case, the leading indicator is captured by the number of private housing units authorized by local building permits, and is obtained from the US Census Bureau. Note, barring the real personal disposable income, all the other variables are available in monthly frequencies. They are converted into their corresponding quarterly values using temporal aggregation, i.e., by taking averages for the monthly values. Note, Figure 2 presents the four census regions of the US economy to provide a general idea about their layout and the census divisions regions and states the four census regions are comprised of.

[INSERT FIGURE 2]

In the VAR models for the different regions, in each equation there are 13 parameters, including the constant, given the fact that the model is estimated with two lags of each variable, with the choice of the lags being confirmed by the unanimity of the sequential modified likelihood ratio (LR) test statistic (each tested at the 5-percent level), the final prediction error (FPE), the Akaike information criterion (AIC) for the Northeast and West regions, FPE and AIC tests for the Midwest, and the FPE criterion and Hannan-Quinn information criterion (HQIC) for the South.[11] While, in the VECMs, we have 10 parameters for Northeast and Midwest and 11 for West and South – these corresponding to the constant, one lag of each of the six variables and the three or four error-correction terms, given three or four cointegrating relationships.[12] The univariate classical (AR) and Bayesian (UVBAR) models are understandably estimated with two lags. All variables, except for the mortgage interest rate and the unemployment rate, have been measured in natural logarithms. Note Sims *et al.* (1990) indicates that with the Bayesian approach entirely based on the likelihood function, the associated inference does not need to take special account of nonstationarity, since the likelihood function has the same Gaussian shape regardless of the presence of nonstationarity. Given this, the variables have been specified in levels.[13]

The 'optimal' Bayesian prior is selected on the basis of Root Mean Squared Error (RMSE) values of the out-of-sample forecasts. Specifically, the six-variable BVARs (BVECMs) are estimated for an initial prior for the period of 1976:Q1 to 2000:Q4 and, then, we forecast 2001:Q1 through 2004:Q3. Since we use two (one) lag(s), the initial two (one) quarter(s) of the sample, 1976:Q1 to 1976:Q2(1), are (is) used to feed the lags. We generate dynamic forecasts, as would naturally be achieved in actual forecasting practice. In each quarter during the forecast period, the models were estimated in order to update the estimate of the coefficient before producing 4-quarters-ahead forecasts. This iterative estimation and 4-step-ahead forecast procedure was carried out for 15 quarters, with the first forecast beginning in 2001:Q1. This experiment produced a total of 15 one-quarter-ahead forecasts, 15-two-quarters ahead forecasts, and so on, up to 15 4-step-ahead forecasts. We use the algorithm in the Econometric Toolbox of MATLAB,[14] for this purpose. The RMSEs[15] for the 15, quarter 1 through quarter 4 forecasts were then calculated for home sales. The average of the MAPE statistic values for one- to four-quarters-ahead forecasts for the period 2001:Q1 to 2004:Q3 are then

examined. We then change the prior and a new set of RMSE values are generated. The combination of the parameter values in the prior, which produces the lowest average RMSE values is selected, as the 'optimal' Bayesian prior. Following Doan (1990) and Dua *et al.* (1999), we choose 0.1 and 0.2 for the overall tightness (w) and 1 and 2 for the harmonic lag decay parameter (d). Moreover, as in Dua and Ray (1995), we also report our results for a combination of $w = 0.3$ and $d = 0.5$. Finally, a symmetric interaction function $f(i, j)$ is assumed with $k_{ij} = 0.7$, as in Dua *et al.* (1999).[16] Note for the univariate BVAR (UBVAR) models, $k_{ij} = 0.001$, which effectively eliminates the 'Vector' part, and allows us to look at autoregression only.

3.2. Evaluation of Forecast Accuracy

In Tables 1 to 4, we compare the RMSEs of one- to four-quarters-ahead out-of-sample-forecasts for the period of 2001:Q1 to 2004:Q3, generated by the RW, AR, VAR, UBVAR, BVAR and BVEC models. At this stage, a few words need to be said regarding the choice of the evaluation criterion for the out-of-sample forecasts generated by the Bayesian models. As Zellner (1986) points out, the 'optimal' Bayesian forecasts will differ depending upon the loss function employed and the form of predictive probability density function. In other words, Bayesian forecasts are sensitive to the choice of the measure used to evaluate the out-of-sample forecast errors. However, Zellner (1986) points out that the use of the mean of the predictive probability density function for a series is optimal relative to a squared error loss function and the Mean Squared Error (MSE), and hence, the RMSE is an appropriate measure to evaluate performance of forecasts. This is exactly what we do below in Tables 1 through 4, when we use the average RMSEs over the one- to four-quarters-ahead forecasting horizon. The conclusions, regarding home sales for each of the four census regions, based on the average one- to four-quarters-ahead RMSEs, from these tables can be summarized as follows:

[INSERT TABLES 1 THROUGH 4]

(i) In three of the four regions, a Bayesian model performed the best among the models compared, specifically, the UBVAR (w=0.2, d=1) for the Northeast, the BVECM (w=0.1, d=2) for the Midwest and the UBVAR (w=0.3, d=0.5) for the West. For the South, the VAR outperformed all other models;

(ii) Focusing first on the South census region, it is interesting to note that the among the Bayesian models, the BVAR, under the different w and d combinations, always outperformed the other Bayesian models, namely, UBVAR and the BVECM models, under the combinations of w and d considered. Further, the average RMSE of all the four BVARs are not very different from the average RMSE of the VAR model, which is the best overall model for the South. Interestingly, barring the VAR and BVAR, all the other models performed almost equally badly irrespective of them being either Bayesian or otherwise. Thus from this analysis, we see that for the South, the model that does best, the VAR, is the one that emphasizes equally the importance of all the lags for all the variables considered in the model to forecast home sales in the South;

(iii) As mentioned above, barring the South, for the remaining three census regions, a Bayesian model always outperformed the other competing models. Specifically, in the Northeast and the West, under all the different w and d combinations, the UBVARs consistently did better than both the BVARs and the BVECMs, while in the Midwest census region, the BVECMs, under all the different w and d combinations, always

outperform the other Bayesian models, while for the West census region, the UBVAR consistently always outperform the other Bayesian models;

(iv)   It is also interesting to point out that for the Northeast, Midwest and the West, the VAR, and among the Bayesian models the BVARs, consistently performed the worst as far as the average RMSE are considered for forecasting.

(v)   Taking (iii) and (iv) into account, we can say that for the Northeast and the West regions, what matters in forecasting home sales is the past values of the variable, with more emphasis being given to the first own lag of home sales, given that the UBVAR models are the standout performers. In addition, the priors imposed for the 'optimal' UBVAR models in these two regions are relatively loose, allowing the own lags to explain most of the variations. With the BVECM models doing the best for the Midwest census region, the result emphasizes the importance of not only short-run dynamics amongst the variables, as it is also the case with the other three regions, but also of the role of long-run equilibrium relationships amongst the variables in predicting home sales.

3.3. Predicting the Recent Downturn in Home Sales of the Four Census Regions
As illustrated in Figure 1, each of housing market in the four census regions experienced a marked reversal of home sales after the peak in 2005:Q3. That is, the home sale peaked and then declined in the four regions. We exposed our classical and 'optimal' Bayesian forecast models to the acid test – of their ability in predicting the peaks and decline in home sales. We estimated the models using data through the third quarter of 2004, and forecasted the home sales for the fourth quarter of 2004. We then updated the data by one quarter, and repeated the forecasting exercise with a model estimated through the fourth quarter of 2004 and forecasted the first quarter of 2005. We continued this updating and forecasting process till the end of the sample in the second quarter of 2009. Table 5 through 8 reports the one-quarter-ahead forecasts for the four census regions.

[INSERT TABLES 5 THROUGH 8]

For the South census region, all the models predicted the peak one quarter immediately after the actual peak of 2005:Q3. For the remaining three census regions, there was at least one model that picked up the peak at, or one quarter ahead, of the actual peak. In the Northeast, the BVECM predicted the peak synchronous to the actual peak in 2005:Q3; while for the VECM it predicted the peak two quarters ahead; and for the remaining RW, UVAR, VAR, UBVAR and BVAR, the peaks were forecast at the immediate quarter following the actual peak. In the Midwest and the West, none of the models considered could pick the peak of 2005:Q3, however, the VECM and BVECM peaked a quarter before in 2005:Q2, while for all the remaining 5 models, the peak was forecast immediately following the actual peak.

4.   Conclusion

This paper analyzes the ability of a random RW, classical and Bayesian versions of AR, VAR and VEC models in forecasting home sales for the four US census regions (Northeast, Midwest, South and West), using quarterly data over the period of 2001:Q1 to 2004:Q3, based on an in-sample of 1976:Q1 till 2000:Q4. We choose the end-point of the out-of-sample horizon as 2004:Q3, since we also use our models to predict the downturn in the home sales of the four census regions (over 2004:Q4 till 2009:Q2) and, hence, stop (a year) prior to the date where the turning point actually

occurred. In our case, the home sales in all the four census regions peaked in 2005:Q3. Following Dua *et al.* (1999), our multivariate forecasting models for each of the four census regions, comprise of home sales, price of homes, mortgage rate, real personal disposable income, unemployment rate and building permits authorized. Understandably, the univariate models only include home sales, as it is the variable of interest.

The main conclusions that could be drawn from this study are as follows: (i) Barring the South, there always exist a specific kind of Bayesian model which tends to outperform all other models in forecasting home sales over the out-of-sample horizon. Specifically, the best performing models are: UBVAR (w=0.2, d=1) for Northeast, BVECM (w=0.1, d=2) for the Midwest, VAR for the South, UBVAR (w=0.3, d=0.5) for the West. This result highlights the fact that economic conditions prevailing at the start of the out-of-sample horizon are not necessarily the same across the regions, and, hence, there does not exist a single model that forecasts best for all the four regions, and; (ii) When we exposed our classical and 'optimal' Bayesian forecast models to predicting the peaks and declines in home sales, we found that barring the South, our models did reasonably well in predicting the turning point exactly at 2005:Q3 or with a lead. For the South, all our models predicted the turning point a quarter after the actual peak. Interestingly, with the exception of the Midwest, we also observed that there is no guarantee that the best performing model over the out-of-sample horizon is also well-suited in predicting the downturn in home sales. In general, the VECMs, both classical and the 'optimal' Bayesian versions, perform the best in predicting the turning point.

The fact that the VECM models, in general, tends to pick up the turning point highlights the importance of modeling both short- and long-run dynamics of the data. The VECM has cointegration relations built into the specification so that it restricts the long-run behavior of the endogenous variables to converge to their cointegrating relationships. The cointegration term, in turn, corrects for the deviation from long-run equilibrium gradually through a series of partial short-run adjustments, and, hence, is perhaps well-suited in capturing the downturn – in other words, deviations from the long-run (equilibrium) relationship. From our forecasting results, one could conclude that as long as the data depicts a specific trend, as has been happening with continuous increases in home sales until 2005:Q3, merely modeling short-run dynamics through a VAR is good enough for forecasting. This is understandable, given that persistence in the data of the variables which, in turn, is captured by the lagged-structure of the VAR model. Further, with the optimal BVECM models performing better than the classical VECM models in terms of forecasting, highlights the importance of the prior restrictions imposed by the Bayesian approach to prevent the over-parameterization problem from creating multicollinearity and larger out-of-sample forecasting errors.

At this stage, it must be pointed out that there are at least two limitations to using the BVAR and BVECM models for forecasting. Firstly, as it is clear from Tables 1 to 4, the accuracy of the forecasts is sensitive to the choice of the priors. Clearly then, if the prior is not well specified, an alternative model used for forecasting may perform better. Secondly, in case of the Bayesian variants, one requires to specify an objective function, for example the RMSEs, to search for the 'optimal' priors, which, in turn, needs to be optimized over the period for which we compute the out-of-sample forecasts. However, there is no guarantee that the chosen parameter values specifying the prior will also be 'optimal' beyond the period for which it was selected, highlighted to some

degree by the turning point exercise. Nevertheless, the ability of the Bayesian models in forecasting and predicting downturns, as shown by our analysis, cannot be underestimated. Given the trend changing pattern of the data, it would be interesting to revisit the issue, in the future, using non-linear and semi- and non-parametric models.

REFERENCES

Carlino, G. A. and DeFina, R. H. The Differential Regional Effects of Monetary Policy. *The Review of Economics and Statistics,* 1998*,* 80, 572–587.

Carlino, G. A. and DeFina, R. H.Do States Respond Differently to Changes in Monetary Policy? Federal Reserve Bank of Philadelphia, *Business Review*, 1999, (July/August), 17–27.

Case, K., Shiller, R. and Quigley, J.Comparing Wealth Effects: The Stock Market

Versus the Housing Market. *Advances in Macroeconomics,* 2005*,* 5, 1-32.

Clapp, J. M. and Giaccotto, C. Evaluating House Price Forecasts**,** *Journal of Real Estate Research*, 2002, 24:1, 1-26.Cooley, T. F. and Leroy, S, F.. Atheoretical Macroeconometrics: A critique. *Journal of Monetary Economics*, Elsevier, 1985,16:3, pages 283-308, November.

Das, S., Gupta, R., and Kabundi, A.. Could We Have Predicted the Recent Downturn in the South African Housing Market? *Journal of Housing Economics*, 2009, 18:4, 325 - 335.

Das, S., Gupta, R., and Kabundi, A. The Blessing of Dimensionality in Forecasting Real House Price Growth in the Nine Census Divisions of the U.S. *Journal of Housing Research*, 2010,19:1,89-109.

Dickey, D. A., Jansen, D. W. and Thornton, D. L. A Primer on Cointegration with an Application to Money and Income. Review. Federal Reserve Bank of St. Louis, 1991, issue Mar,58-78.

Doan, T.A. Regression Analysis of Time Series. VAR Econometrics. Evanston, Il, 1990.

Doan, T. A., Litterman, R. B. and Sims, C. A. Forecasting and Conditional Projections   Using Realistic Prior Distributions. *Econometric Reviews*, 1984, 3:1, 1-100.

Dua, P. and Ray, S. C. A BVAR Model for the Connecticut Economy. *Journal of   Forecasting,* 1995, 14:3, 167-180.

Dua, P. and Smyth, D.J. Forecasting U.S. Home Sales Using BVAR Models and Survey

Data on Households' Buying Attitudes for Homes. *Journal of Forecasting*, 1995, 14, 167-180.

Dua, P. and Miller, S. M. Forecasting Connecticut Home Sales in a BVAR Framework   Using Coincident and Leading Indexes. *Journal of Real Estate Finance and Economics*, 1996, 13:3,   219-235.

Dua, P., Miller, S. M. and Smyth, D. J. Using Leading Indicators to Forecast US home sales   in Bayesian VAR framework. *Journal of Real Estate Finance and Economics*, 1999, 18:2,   191-205.

Ghent, A. Sticky Mortgages and the Real Effects of Monetary Policy. Mimeo, Department of Real Estate**,** Zicklin School of Business, Baruch College, CUNY, 2009.

Ghent, A., and Owyang, M. Is Housing the Business Cycle? Evidence from US Cities. Mimeo, Department of Real Estate**,** Zicklin School of Business, Baruch College, CUNY, 2009.

Green, R. Follow the Leader: How Changes in Residential and Non-Residential Investment Predict Changes in GDP. *Real Estate Economics,* 1997*, 25*, 253–270.

Gupta, R. and Sichei, M. M. A BVAR Model for the South African Economy. *South African   Journal of Economics*, 2006, 74:3, 391-409.

Gupta, R. Forecasting the South African Economy with VARs and VECMs. *South African Journal of Economics,* 2006, 74:4, 611-628.

Gupta, R., Jurgilas, M. and Kabundi, A. The Effect of Monetary Policy on House Price Inflation: A Factor Augmented Vector Autoregression (FAVAR) Approach. *Journal of Economic Studies,* Forthcoming.

Hafer, R. W. and Sheehan, R. G. The Sensitivity of VAR forecasts to alternative lag structures. *International Journal of Forecasting*, 1989, Elsevier,5:3, 399-408.

Hamilton, J. D. Time Series Analysis. Princeton University Press, Princeton, N. J, 1994.

Iacoviello, M. House Prices, Borrowing Constraints, and Monetary Policy in th   Business Cycle. *American Economic Review, 2005, 95*, 739–764.

Iacoviello, M., and Neri S. Housing Market Spillovers: Evidence from an Estimated DSGE Model. Department of Economics, Working Paper No. 659, Boston College, 2008.

Iacoviello, M.and Neri, S. Housing Market Spillovers: Evidence from an Estimated DSGE Model. *American Economic Journal: Macroeconomics*, in press, 2010.

Johansen, S. Likelihood-Based Inference in Cointegrated Vector Auto-regressive Models. Oxford University Press, Inc., New York, 1995.

Leamer, E. E. Housing is the business cycle. Working Paper 13428, National Bureau of Economic Research, 2007.

LeSage, J. P. Applied Econometrics Using MATLAB. www.spatial-econometrics.com, 1999.

LeSage, J. P. A Comparison of the Forecasting Ability of ECM and VAR Models. *The Review of Economics and Statistics*, 1990, MIT Press, 72:4, pages 664-71, November.

Litterman, R. B. A Bayesian Procedure for Forecasting with Vector Autoregressions. Working Paper, Federal Reserve Bank of Minneapolis, 1981.

Litterman, R. B. Forecasting with Bayesian Vector Autoregressions – Five Years of   Experience. *Journal of Business and Economic Statistics*, 1986, 4:1, 25-38.

Sims, C. A., Stock, J. H. and Watson, M. W. Inference in Linear Time Series Models with Some Unit Roots. *Econometrica*, Econometric Society, 1990, 58:1, 113-44, January.

Sims, C. A. Macroeconomics and Reality. *Econometrica*, 1980, 48:1, 1-48.

Song, Shi, Young M. and Hargreaves, B. House Price - Volume Dynamics: Evidence from 12 Cities in New Zealand, *Journal of Real Estate Research*, 2010, 32:1, 75- 99.

Spencer, D. E. Developing a Bayesian Vector Autoregression Model. *International   Journal of Forecasting*, 1993, 9:3, 407-421.

Theil, H. Principles of Econometrics. New York: John Wiley, 1971.

Todd, R. M. Improving Economic Forecasting with Bayesian Vector Autoregression. *Quarterly Review*, 1984, Federal Reserve Bank of Minneapolis, Fall, 18-29.

Vargas-Silva, C. Monetary Policy and the US Housing Market: A VAR Analysis Imposing Sign Restrictions. *Journal of Macroeconomics*, 2008a, *30*, 977–990.

Vargas-Silva, C. The Effect of Monetary Policy on Housing: A Factor-Augmented   Vector Autoregression (FAVAR) Approach. *Applied Economics Letters,* 2008b, 15, 749 -752.

Zellner, A. 'A Tale of Forecasting 1001 Series: The Bayesian Knight Strikes Again'. *International Journal of Forecasting,* 1986, 2:494-494.

Zellner, A. Statistical analysis of econometric models. *Journal of the American Satistical Association*, 1979, 74, 628–643.

Zellner, A. and Palm, F. Time Series Analysis and Simultaneous Equation Econometric   Models. *Journal of Econometrics*, 1974, 2, 17-54.
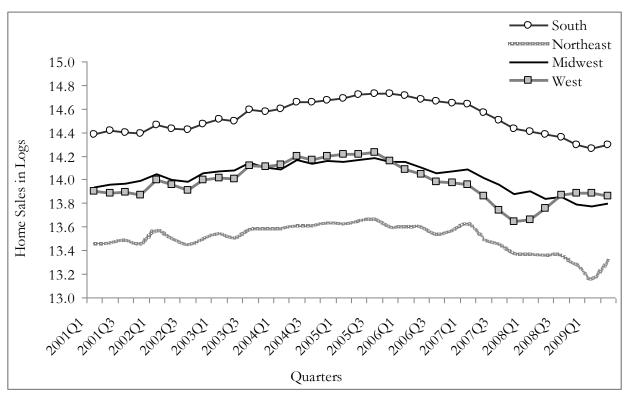
Figure 1: The Recent Downturn in Home Sales of the Four US Census Regions (2001:Q1-2009:Q2)
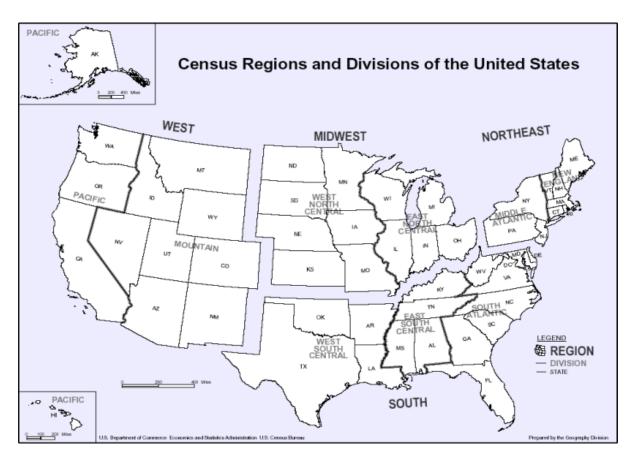
.

Figure 2: Layout of the Four Census Regions of the US.

(Source: US Department of Commerce, Economics and Statistics Administration, US Census Bureau.)[17]

TABLE 1:   One- to Four-Quarters-Ahead (QA) RMSEs for South (2001:Q1-2004:Q3)

Parameterization	Models	QA 1	2	3	4	Average
	RW	0.8219	1.8884	1.9085	1.9148	1.6334
	AR	0.8791	2.0482	2.1601	2.2475	1.8337
	VAR	0.4293	0.2887	0.4402	0.7825	**0.4852**
	VECM	1.0709	1.1576	2.4660	3.6935	2.0970
	UBVAR	0.7030	1.7125	1.6795	1.6295	1.4311
w=0.3 d=0.5	BVAR	0.4209	0.2995	0.4475	0.7863	0.4886
	BVECM	1.0562	1.1380	2.4424	3.6651	2.0754
	UBVAR	0.7266	1.7619	1.7555	1.7325	1.4941
w=0.2 d=1	BVAR	0.3996	0.3246	0.4626	0.7913	0.4945
	BVECM	1.0273	1.1010	2.3991	3.6142	2.0354
	UBVAR	0.7663	1.8436	1.8798	1.9002	1.5975
w=0.1 d=1	BVAR	0.3253	0.4123	0.5187	0.8138	0.5175
	BVECM	0.9216	0.9622	2.2325	3.4147	1.8828
	UBVAR	0.7304	1.7679	1.7634	1.7428	1.5011
w=0.2 d=2	BVAR	0.3428	0.3864	0.4961	0.7948	0.5050
	BVECM	0.9726	1.0369	2.3289	3.5366	1.9688
	UBVAR	0.7624	1.8288	1.8534	1.8634	1.5770
w=0.1 d=2	BVAR	0.1891	0.5534	0.5989	0.8228	0.5411
	BVECM	0.8127	0.8377	2.0948	3.2639	1.7523

Note: The RMSEs are in percentages.

TABLE 2: One- to Four-Quarters-Ahead (QA) RMSEs for Northeast (2001:Q1-2004:Q3)

Parameterization	Models	QA 1	2	3	4	Average
	RW	0.2375	0.2324	0.5338	0.5193	0.3808
	AR	0.2442	0.2419	0.5144	0.5542	0.3887
	VAR	1.7014	1.9674	1.1375	2.1106	1.7292
	VECM	1.5063	0.8883	0.0823	0.5199	0.7492
	UBVAR	0.1516	0.0809	0.7510	0.2362	0.3049
w=0.3 d=0.5	BVAR	1.6956	1.9660	1.1388	2.1120	1.7281
	BVECM	1.5049	0.8880	0.0753	0.5264	0.7487
	UBVAR	0.0976	0.0332	0.9252	0.0025	**0.2646**
w=0.2 d=1	BVAR	1.6848	1.9643	1.1425	2.1165	1.7270
	BVECM	1.5058	0.8928	0.0583	0.5440	0.7502
	UBVAR	0.0469	0.1416	1.0921	0.2235	0.3760
w=0.1 d=1	BVAR	1.6356	1.9467	1.1470	2.1248	1.7135
	BVECM	1.5056	0.9027	0.0036	0.5948	0.7517
	UBVAR	0.0898	0.0504	0.9519	0.0337	0.2815
w=0.2 d=2	BVAR	1.6646	1.9631	1.1509	2.1283	1.7267
	BVECM	1.5209	0.9207	0.0172	0.5915	0.7626
	UBVAR	0.0450	0.1470	1.1013	0.2366	0.3825
w=0.1 d=2	BVAR	1.5667	1.9093	1.1317	2.1220	1.6824
	BVECM	1.5566	0.9741	0.0789	0.6860	0.8239

Note: The RMSEs are in percentages.

TABLE 3:   One- to Four-Quarters-Ahead (QA) RMSEs for Midwest (2001:Q1-2004:Q3)

Parameterization	Models	QA 1	2	3	4	Average
	RW	0.5184	1.5295	1.7441	2.7128	1.6262
	AR	0.6203	1.7301	2.0168	3.0352	1.8506
	VAR	0.0105	1.6256	2.3524	3.6987	1.9218
	VECM	0.5585	0.1916	0.5407	0.1376	0.3571
	UBVAR	0.4307	1.3916	1.5643	2.4943	1.4702
w=0.3 d=0.5	**BVAR**	0.0067	1.6346	2.3599	3.7022	1.9259
	BVECM	0.5441	0.1792	0.5304	0.1296	0.3458
	UBVAR	0.4774	1.4913	1.7159	2.6958	1.5951
w=0.2 d=1	**BVAR**	0.0596	1.6651	2.3801	3.7094	1.9536
	BVECM	0.5152	0.1561	0.5127	0.1178	0.3255
	UBVAR	0.5340	1.6105	1.8963	2.9350	1.7440
w=0.1 d=1	**BVAR**	0.2172	1.7615	2.4421	3.7344	2.0388
	BVECM	0.4117	0.0690	0.4419	0.0649	0.2469
	UBVAR	0.4846	1.5054	1.7367	2.7232	1.6125
w=0.2 d=2	**BVAR**	0.2200	1.7652	2.4378	3.7268	2.0375
	BVECM	0.4598	0.1172	0.4880	0.1090	0.2935
	UBVAR	0.5315	1.6009	1.8796	2.9123	1.7311
w=0.1 d=2	**BVAR**	0.5463	1.9972	2.5728	3.7784	2.2237
	BVECM	0.3062	0.0056	0.3959	0.0503	**0.1895**

Note: The RMSEs are in percentages.

TABLE 4:   **One- to Four-Quarters-Ahead (QA) RMSEs for West (2001:Q1-2004:Q3)**

Parameterization	Models	QA 1	2	3	4	Average
	RW	0.5541	0.8131	1.5513	1.2195	1.0345
	AR	0.6862	1.2303	2.2287	2.1159	1.5653
	VAR	0.0011	0.9248	2.4794	2.8129	1.5546
	VECM	1.1005	1.4622	1.2587	2.0858	1.4768
	UBVAR	0.2324	0.3361	0.9528	0.5069	**0.5071**
w=0.3 d=0.5	BVAR	0.0108	0.9342	2.4875	2.8220	1.5636
	BVECM	1.0883	1.4499	1.2486	2.0771	1.4660
	UBVAR	0.2534	0.3831	1.0270	0.6088	0.5681
w=0.2 d=1	BVAR	0.0388	0.9631	2.5143	2.8523	1.5921
	BVECM	1.0616	1.4236	1.2272	2.0593	1.4429
	UBVAR	0.3533	0.5922	1.3442	1.0323	0.8305
w=0.1 d=1	BVAR	0.1350	1.0623	2.6040	2.9515	1.6882
	BVECM	0.9681	1.3316	1.1537	1.9992	1.3632
	UBVAR	0.2862	0.4471	1.1208	0.7322	0.6466
w=0.2 d=2	BVAR	0.1236	1.0561	2.6046	2.9541	1.6846
	BVECM	1.0037	1.3682	1.1838	2.0248	1.3951
	UBVAR	0.3537	0.5759	1.3074	0.9766	0.8034
w=0.1 d=2	BVAR	0.3512	1.3060	2.8437	3.2152	1.9290
	BVECM	0.8586	1.2328	1.0845	1.9542	1.2825

Note: The RMSEs are in percentages.

Table 5: Recursive Forecasts: 2004:Q4 to 2009:Q2 for South

Quarters	ACTUAL	RW	UVAR	VAR	VECM	UBVAR	BVAR	BVECM
2004Q3	14.6513	14.6513	14.6513	14.6513	14.6513	14.6513	14.6513	14.6513
2004Q4	14.6742	14.6559	14.6564	14.6646	14.6691	14.6501	14.6644	14.6673
2005Q1	14.6882	14.6774	14.6786	14.67	14.6842	14.6741	14.6699	14.6813
2005Q2	14.7170	14.6964	14.6989	14.6836	14.6942	14.6948	14.6836	14.6915
2005Q3	**14.7278**	14.7276	14.7291	14.718	14.7228	14.7268	14.718	14.7234
2005Q4	14.7265	**14.7381**	**14.7404**	**14.7636**	**14.7653**	**14.735**	**14.7634**	**14.7626**
2006Q1	14.7116	14.7362	14.7357	14.702	14.716	14.734	14.7021	14.716
2006Q2	14.6812	14.7137	14.7131	14.7099	14.7237	14.7115	14.7098	14.7229
2006Q3	14.6599	14.6805	14.6792	14.6645	14.6802	14.6796	14.6646	14.678
2006Q4	14.6455	14.6592	14.6575	14.6405	14.654	14.6602	14.6405	14.6541
2007Q1	14.6411	14.6505	14.6483	14.6411	14.6538	14.6513	14.6409	14.6534
2007Q2	14.5685	14.6404	14.6388	14.6187	14.6377	14.6414	14.6187	14.6348
2007Q3	14.5020	14.5695	14.5656	14.5543	14.5669	14.5644	14.5542	14.5658
2007Q4	14.4271	14.4984	14.4931	14.4562	14.4745	14.4938	14.4563	14.4719
2008Q1	14.4088	14.4309	14.4247	14.4059	14.4149	14.4238	14.4059	14.4147
2008Q2	14.3789	14.4124	14.4104	14.3771	14.387	14.4112	14.3771	14.385
2008Q3	14.3540	14.3835	14.3814	14.3441	14.3296	14.3803	14.3441	14.3298
2008Q4	14.2897	14.3529	14.3523	14.2977	14.2939	14.3505	14.2979	14.2928
2009Q1	14.2645	14.2926	14.2877	14.1951	14.2144	14.2866	14.1953	14.216
2009Q2	14.2917	14.2664	14.2615	14.1947	14.1842	14.2615	14.1947	14.1921

Note: Bayesian models corresponds to the optimal ones; Bold numbers equal the maximum values in each column.

Table 6: Recursive Forecasts: 2004:Q4 to 2009:Q2 for Northeast

Quarters	ACTUAL	RW	UVAR	VAR	VECM	UBVAR	BVAR	BVECM
2004Q3	13.6089	13.6089	13.6089	13.6089	13.6089	13.6089	13.6089	13.6089
2004Q4	13.6372	13.6024	13.6021	13.5499	13.5984	13.6033	13.5456	13.5985
2005Q1	13.6292	13.6308	13.6289	13.5737	**13.6442**	13.6323	13.5725	13.644
2005Q2	13.6530	13.6206	13.6206	13.5671	13.6306	13.6236	13.5648	13.6305
2005Q3	**13.6647**	13.6399	13.6368	13.5735	13.6366	13.6428	13.5745	**13.6366**
2005Q4	13.6048	**13.6501**	**13.6476**	**13.5895**	13.6366	**13.6533**	**13.5893**	13.6365
2006Q1	13.6048	13.5956	13.5952	13.5402	13.5896	13.5967	13.5442	13.5894
2006Q2	13.6007	13.5972	13.5948	13.5675	13.6047	13.5961	13.5676	13.6047
2006Q3	13.5411	13.5938	13.5927	13.5358	13.57	13.5965	13.534	13.57
2006Q4	13.5713	13.5413	13.539	13.5123	13.5331	13.5401	13.5082	13.5331
2007Q1	13.6252	13.5688	13.5649	13.5563	13.5924	13.5675	13.5509	13.5924
2007Q2	13.4962	13.6227	13.6176	13.577	13.615	13.6249	13.5759	13.6153
2007Q3	13.4588	13.5071	13.509	13.5004	13.5312	13.5058	13.4957	13.5312
2007Q4	13.3744	13.4722	13.4688	13.4579	13.5045	13.4647	13.4602	13.5043
2008Q1	13.3692	13.3947	13.3928	13.4436	13.4366	13.3794	13.4412	13.4365
2008Q2	13.3640	13.392	13.3912	13.4409	13.4235	13.3787	13.4369	13.4238
2008Q3	13.3640	13.3894	13.3843	13.3976	13.3565	13.3693	13.4133	13.3568
2008Q4	13.2708	13.3824	13.383	13.37	13.3295	13.3693	13.368	13.3299
2009Q1	13.1616	13.3003	13.2993	13.3009	13.2663	13.2839	13.2949	13.2667
2009Q2	13.3047	13.1981	13.1964	13.2203	13.2331	13.1753	13.2202	13.2328

Note: See notes to Table 5.

Table 7: Recursive Forecasts: 2004:Q4 to 2009:Q2 for Midwest

Quarters	ACTUAL	RW	UVAR	VAR	VECM	UBVAR	BVAR	BVECM
2004Q3	14.1352	14.1352	14.1352	14.1352	14.1352	14.1352	14.1352	14.1352
2004Q4	14.1544	14.1319	14.1252	14.0987	14.0925	14.1296	14.0987	14.0952
2005Q1	14.1448	14.1422	14.1398	14.1161	14.146	14.1445	14.1162	14.1371
2005Q2	14.1662	14.1319	14.1293	14.141	**14.1644**	14.1319	14.141	**14.1624**
2005Q3	**14.1778**	14.161	14.1619	14.1493	14.1351	14.1675	14.1495	14.1382
2005Q4	14.1472	**14.1711**	**14.1692**	**14.1552**	14.1577	**14.1748**	**14.1553**	14.1592
2006Q1	14.1520	14.1408	14.1354	14.1042	14.1288	14.1399	14.1045	14.1262
2006Q2	14.1032	14.1408	14.1379	14.1221	14.1443	14.1447	14.1221	14.143
2006Q3	14.0519	14.0906	14.0835	14.059	14.0928	14.089	14.0591	14.0909
2006Q4	14.0650	14.0432	14.036	14.0324	14.0502	14.0406	14.0323	14.0519
2007Q1	14.0881	14.0531	14.0524	14.0513	14.0759	14.0595	14.0513	14.0763
2007Q2	14.0144	14.0814	14.0818	14.0327	14.0617	14.0917	14.033	14.0627
2007Q3	13.9524	14.0062	13.9959	13.9704	13.9967	13.9987	13.9705	13.9946
2007Q4	13.8738	13.9493	13.9425	13.9166	13.9379	13.9437	13.9168	13.9353
2008Q1	13.9017	13.8759	13.8667	13.8834	13.8887	13.8633	13.8833	13.8886
2008Q2	13.8353	13.9055	13.908	13.8592	13.8944	13.9101	13.8594	13.8924
2008Q3	13.8483	13.8522	13.8444	13.7324	13.7456	13.8391	13.7331	13.7451
2008Q4	13.7885	13.861	13.8632	13.7867	13.7522	13.8598	13.7872	13.7606
2009Q1	13.7747	13.805	13.7987	13.759	13.7519	13.7915	13.7589	13.7546
2009Q2	13.7987	13.7913	13.7872	13.6982	13.7567	13.7781	13.6984	13.7497

Note: See notes to Table 5.

Table 8: Recursive Forecasts: 2004:Q4 to 2009:Q2 for West

Quarters	ACTUAL	RW	UVAR	VAR	VECM	UBVAR	BVAR	BVECM
2004Q3	14.1662	14.1662	14.1662	14.1662	14.1662	14.1662	14.1662	14.1662
2004Q4	14.1939	14.1564	14.1527	14.1879	14.1966	14.1566	14.1875	14.1964
2005Q1	14.2165	14.1779	14.1771	14.1998	14.2224	14.1857	14.1995	14.2211
2005Q2	14.2120	14.2051	14.2073	14.2202	**14.2565**	14.2177	14.2199	**14.2482**
2005Q3	**14.2276**	14.1939	14.1917	14.2101	14.194	14.2015	14.21	14.1989
2005Q4	14.1591	**14.2145**	**14.214**	**14.223**	14.2146	**14.227**	**14.223**	14.2192
2006Q1	14.0830	14.1394	14.1318	14.1296	14.1524	14.1431	14.1298	14.1475
2006Q2	14.0466	14.0582	14.0472	14.0768	14.1027	14.0622	14.0766	14.1019
2006Q3	13.9838	14.0281	14.0213	14.0143	14.0807	14.0416	14.0143	14.0707
2006Q4	13.9754	13.9569	13.9471	13.9679	13.9751	13.9664	13.9675	13.9769
2007Q1	13.9553	13.9585	13.9545	13.9758	14.015	13.9788	13.9755	14.0082
2007Q2	13.8579	13.9425	13.9359	13.933	13.9353	13.9563	13.9331	13.9363
2007Q3	13.7357	13.8427	13.8317	13.8378	13.8649	13.8453	13.8377	13.8583
2007Q4	13.6412	13.7264	13.7123	13.7034	13.7255	13.724	13.7036	13.7234
2008Q1	13.6569	13.6323	13.6196	13.6434	13.6633	13.6298	13.6432	13.6609
2008Q2	13.7536	13.6602	13.6583	13.6647	13.6995	13.6738	13.6642	13.6956
2008Q3	13.8643	13.758	13.7641	13.7084	13.7424	13.776	13.7087	13.7396
2008Q4	13.8863	13.8743	13.8852	13.8694	13.8676	13.8882	13.8692	13.8741
2009Q1	13.8832	13.914	13.9185	13.9688	13.9953	13.901	13.9681	13.9924
2009Q2	13.8579	13.9204	13.9206	14.0004	14.075	13.8835	13.9999	14.0567

Note: See notes to Table 5.

ENDNOTES

[1] Besides the home sales literature, our choice of variables is also in line with variables generally used to forecast house prices using autoregressive models. In this regard, see for example Clapp and Giaccotto (2002) and Song *et al.* (2010).
[2] It is likely that home sales would also portray heterogeneity in behavior depending on the size of houses we are looking at (Das *et al.,* 2009). Data unavailability, however, precludes us from investigating this line of thinking.
[3] As in Dua *et al.* (1999), replacing housing permits authorized by housing starts, as a measure of leading indicator, failed to improve the predictive ability of our models. These results are, however, available upon request from the authors.
[4] The discussion in this section relies heavily on LeSage (1999), Gupta and Sichei (2006), Gupta (2006) and Das *et al.* (2010).
[5] $A(L) = A_1 L + A_2 L^2 + ... + A_p L^p$; and A_0 equals an $(n \times 1)$ vector of constant terms.
[6] A series is integrated of order *q,* if it requires *q* differences to transform it into a zero-mean, purely non-deterministic stationary process.
[7] See LeSage (1990) and references cited therein for further details regarding the non-stationarity of most macroeconomic time series.
[8] See Dickey *et al.* (1991) and Johansen (1995) for further technical details.
[9] For an illustration, see Dua and Ray (1995).
[10] Data on home prices were not available in their seasonally adjusted form, and hence, were deseasonalized using the Census X11 (multiplicative) method – the standard procedure used by the US Census Bureau to seasonally adjust publicly released data.
[11] Hafer and Sheehan (1989) find that the accuracy of the forecasts from the VAR is sensitive to the choice of lags. Their results indicated that shorter-lagged models are more accurate, in terms of forecasts, than longer lag models.
[12] The cointegrating relationships are based on the trace statistics and Maximum Eigen-value statistic compared to the critical values at the 95 percent level.
[13] However, using the Augmented Dickey Fuller (ADF), the Phillips-Perron (PP), the GLS-detrended Dickey-Fuller (DF-GLS) and the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) tests all the 6 variables were found to be, first-order difference stationary, i.e., integrated of order 1 *(I(1))*.

[14] All statistical analysis was performed using MATLAB, version R2009a.

[15] Note that if A_{t+n} denotes the actual value of a specific variable in period $t + n$ and $_t F_{t+n}$ is the forecast made in period t for $t + n$, the RMSE statistic can be defined as: $\sqrt{\dfrac{\sum (A_{t+n} - _t F_{t+n})^2}{N}}$. For $n = 1$, the summation runs from 2001:Q1 to 2004:Q3, and for $n = 2$, the same covers the period of 2001:Q2 to 2004:Q3 and so on.

[16] In the standard Minnesota prior $k_{ij} = 0.5$. However, this value deteriorated our BVAR and BVECM forecasts when compared to 0.7.

[17] See: http://commons.wikimedia.org/wiki/File:Census_Regions_and_Divisions.PNG.