Goodness C. AYE, PhD
Email: goodness.aye@gmail.com
Department of Economics
University of Pretoria, South Africa
Professor Rangan GUPTA, PhD
Email: rangan.gupta@up.ac.za
Department of Economics
University of Pretoria, South Africa

FORECASTING REAL HOUSE PRICE OF THE U.S.: AN ANALYSIS COVERING 1890 TO 2012

Abstract. *This paper evaluates the ability of Bayesian shrinkage-based dynamic predictive regression models estimated with hierarchical priors (Adaptive Jefferys, Adaptive Student-t, Lasso, Fussed Lasso and Elastic Net priors) and non-hierarchical priors (Gaussian, Lasso-Lars, Lasso-Landweber) in forecasting the U.S. real house price growth. We also compare results with forecasts from bivariate OLS regressions and principal component regression. We use annual dataset on 10 macroeconomic predictors spanning the period 1890 to 2012. Using an out-of-sample period of 1917 to 2012, our results based on MSE and* McCracken (2007) *MSE-F statistic, indicate that in general, the non-hierarchical Bayesian shrinkage estimators perform better than their hierarchical counterparts as well as the least square estimators. The Bayesian shrinkage estimated with Lasso-Landweber is the best-suited model for forecasting the U.S. real house price. Among the least square models, the individual regression with house price regressed on the fiscal policy variable outperforms the rest. Also results from Lasso-Landweber portray the fiscal policy variable as the best predictor of the U.S. house prices especially in the recent times while the short-term interest rate and real construction cost also did well at the beginning and middle of the sample.*
Keywords*: Real house price, forecasting, predictive, shrinkage, hierarchical, non-hierarchical, least squares*

JEL Classification: C32, C53, R31

1. Introduction

The recent global financial crisis has resuscitated interest in both research and policy circles as to the role of asset prices in general and house price in particular in the economy. This paper examines the ability of dynamic predictive regression models in forecasting house price for the U.S. economy. We consider Bayesian non-hierarchical shrinkage priors, Bayesian hierarchical shrinkage priors, ordinary least square regressions and principal component regression. We use time series data on 10 macroeconomic predictors. These are fiscal policy, real GDP per capita, unemployment, long term interest rate, short term interest rate, inflation rate, population, real construction cost, real stock price and real oil price. With all these in place, three key questions arise. First, why is it important to forecast house price growth in general and the U.S. house price growth in particular? Second, which predictor(s) are more important in forecasting real house price growth for the U.S.? Third and related to the second, why do we need to consider so many shrinkage forecasting models?

As far as the answer to the first question is concerned, a strong motivation can be found in Leamer (2007), who argues that "Housing is the Business Cycle" (p. 149). He performs a battery of empirical analyses on the business cycle, which he calls the consumer cycle because of the importance of residential investment and durable consumption spending in explaining the onset of recessions. Excluding the most recent Great Recession that he did not consider, residential investment and durable consumption experienced significant problems before the beginning of eight of the ten post World War II recessions. Leamer (2007) argues that the characteristic of the housing market make it a crucial factor in explaining recessions. To wit, the stock-flow nature of the housing market and the reluctance of home owners to lower their prices in a weak market provide the setting for cyclical movement in sales volume. And the cyclical movement in sales volume implies cyclical movements in housing construction and employment. When the economy booms, construction and employment in the housing sector expand, along with increases in nominal house price. During the contraction, nominal house price fall sluggishly and most of the adjustment arises through decreases in sales volume and, thus, construction and employment activity in housing. The Great Recession proved the exception as nominal house price dropped dramatically around the country. Although nominal house price typically fall sluggishly, real house price does fall during recessions as general inflationary trends reduce real house price even with sticky nominal house price. Our analysis focuses on forecasting real house price growth. In sum, during a boom period, developers overbuild the supply of new housing. The size of the excess building, which depends on the strength and length of the boom, will help to determine the length of the next recession. Good monetary policy requires action before the overbuilding goes too far and necessitates central bank intervention early in the boom period, when political pressure probably weighs against monetary policy restraint. That is, understanding and forecasting movements in the housing market plays a critical role for monetary policy authorities.

Further, more recently, several authors argue that asset prices help forecast both inflation and output (Stock and Watson, 2003; Gupta and Das, 2010; Das et al., 2010 among others). Since homes imbed much individual wealth, house price movements may provide important signals for consumption, output, and inflation. That is, housing market adjustments play an important role in the business cycle (Iacoviello and Neri, 2010), not only because housing investment proves a volatile component of demand, but also because house price changes generate important wealth effects on consumption and investment. Leamer (2007) states an even stronger case, as we noted above, arguing that housing is the business cycle. In sum, models that forecast real house price can give policy makers and other stake holders an idea about the future direction of the economy, and hence, can provide important information for designing better and more-appropriate policies. Leamer (2007) notes that the housing market predicted 8 of the 10 post World War II recessions. If he wrote his paper today, the analysis probably would argue that the housing market predicted 9 of the 11 post World War II recessions. In other words, the housing sector acts as a leading indicator for the real sector of the economy. The recent world-wide credit crunch began with the burst of the house-price bubble, which, in turn, led the real sector of the world's economy toward an economic slump. Therefore, predicting real house price correctly cannot be overemphasized.

With respect to the second question, Korobilis (2013a) noted the importance of shrinkage of the predictors: First, even when all predictors are relevant and the full model with all predictors included is the correct (unbiased) model, it is still possible to find a biased model with a lower in-sample mean square error (MSE). However, one can achieve a much lower variance of the coefficient estimates and lower the MSE by introducing some bias in the form of shrinking some of the coefficients in the full model. Second, heavily parameterised models tend to be over-fitted in-sample, and provide very poor out-of-sample fits. Again, it is possible to achieve parsimony and enhance the economic interpretation of results, by introducing some sort of penalty for very complex models through shrinkage and focusing on a model with a few useful predictors. Our model consists of ten predictors which are selected based on standard literature. For recent studies that use GDP per capita (Agnello and Schuknecht, 2011); interest rate (Mikhed and Zemčík, 2009; Agnello and Schuknecht, 2011); population (Agnello and Schuknecht, 2011; Mikhed and Zemčík, 2009); stock price (Mikhed and Zemčík, 2009; Rapach and Strauss, 2009); construction cost (Mikhed and Zemčík, 2009; Zeno and Füss, 2010); unemployment/employment (Rapach and Strauss, 2007); inflation (Rapach and Strauss, 2007); Fiscal policy (Afonso and Sousa, 2011, 2012; Agnello and Schuknecht, 2011), oil price (Beltratti and Morana, 2010).

The answer to the third question lies in the difficulty in forecasting economic variables such as the real house price, since the forecast depends on the models used in generating them. Efficiency in using information from many predictors for improving forecast accuracy is very important. It is now increasingly realized that models that use more information aside that from the real house price itself is likely to improve the forecast over the models that do not use such information. There are two broad approaches for incorporating information from a large number of data series – extracting common factors, i.e., principle components (Stock and Watson 2002a,b and Forni et al., 2005 cited in De Mol et al., 2008) and Bayesian shrinkage methods (De Mol et al., 2008; Korobilis, 2013a, 2013b). Therefore it is crucial to evaluate forecasts from different models and select the 'best' based on an objective criterion (Dua et al., 2008). A model is considered superior to its competitors if it produces smaller forecast errors than its competitors. On this basis, we evaluate the forecasts from different shrinkage models as enumerated above using the mean square error (MSE) relative to an autoregressive benchmark model, and we test for the significance of the MSEs using the McCracken (2007) MSE-F statistic.

To realize the contribution of this study, it is important to place this paper in the context of current research that focusing on forecasting in the housing market. In this regard, few studies are worth mentioning: Rapach and Strauss (2007) used an autoregressive distributed lag (ARDL) model framework, containing 25 determinants, to forecast real housing price growth for the individual states of the Federal Reserve's Eighth District. Given the difficulty in determining a priori particular variables that are most important for forecasting real housing price growth, the authors also use various methods to combine the individual ARDL model forecasts, which result in better forecast of real housing price growth. Rapach and Strauss (2009) perform the same analysis for the 20 largest U.S. states based on ARDL models containing large number of potential predictors, including state, regional and national level variables. Once again, the authors reach similar conclusions on the importance of combining forecasts. On the other hand, Gupta and Das (2010), look into forecasting the recent downturn in real house price growth rates for the twenty largest states of the U.S. economy. In this paper, the authors use Spatial Bayesian VARs (BVARs), based merely on real house price growth rates, to predict their downturn over the period of 2007:01 to 2008:01. They find that, though the models are quite well-equipped in predicting the recent downturn, they underestimate the decline in the real house price growth rates by quite a margin. They attribute this under-prediction of the models to the lack of any information on fundamentals in the estimation process. Gupta and Kabundi (2010) used Bayesian and principal component regressions which did not allow for lags of the variables (both predictors and the national house price), and hence, suffer from possible problems of endogeneity. Das et al., (2010) used small-scale BVARs, BFAVARs and large-scale BVARs to solve this problem in forecasting house prices of the nine census regions. The latter study used the standard Minnesota Bayesian prior in estimating the Bayesian models. Gupta et al. (2011) examine the explanatory power of including information from a large set of economic variables, using VAR, FAVAR, and various Bayesian time-series models- small and large scale BVAR, and BFAVAR. Based on the average MSE for the one-, two-, three-, and four-quarters-ahead forecasts, they find that the small-scale Bayesian-shrinkage model (10 variables) outperforms the other models, as well as outperforming the large-scale Bayesian-shrinkage model. Also Gupta (2013) using dynamic factor and Bayesian shrinkage models and large number of predictors (145 variables) forecasts house prices for four U.S. census regions and the aggregate economy. The results show that the factor-based models were in general the best.

The studies involving Bayesian methods have basically compared the performance of non-hierarchical priors and other models. Given the increasing popularity of hierarchical Bayes priors (Kyung, et al., 2010 cited in Korobilis, 2013a; Korobilis, 2013a), we extend the previous studies by examining the forecasting performance of non-hierarchical Bayesian and hierarchical Bayesian shrinkage priors. We also compare results with least square and principal component regressions. More importantly, unlike previous studies, we also contribute by using a much longer sample (1890-2012). For instance, Rapach and Strauss (2007, 2009) used samples covering 1975-2005, Das et al. (2010) and Gupta and Kabundi (2010a), use samples covering 1991 till mid-2005; Gupta et al. (2011) and Gupta (2013) use samples covering 1976 to 2005 and 1968 to 2012, respectively. Using longer sample period allows us to look at the forecasting performances of these models over periods of drastic and prolonged changes in the trends of house price. The rest of the paper is organized as follows: the methodology is discussed in section 2. Data and empirical results are presented in section 3 while section 4 concludes.

2. Methodology

We employ a number of dynamic predictive regression models for forecasting the U.S. real house price. These include the ordinary least square predictive regression using all the 10 predictors that we consider, and 10 individual least square predictive regressions, principal component regression, Bayesian predictive regressions estimated with five hierarchical shrinkage (adaptive shrinkage Jeffreys, adaptive shrinkage student-t, hierarchical lasso, hierarchical fussed lasso, hierarchical elastic net) priors as in Korobilis (2013a), three non-hierarchical Bayesian priors (Gaussian prior, least absolute shrinkage and selection operator with least angle regression algorithm, lasso with iterative Landweber algorithm) as in De Mol et al. (2008). We discuss these forecasting models in turn.

The starting point is a dynamic regression of the form:

$$y_{t+h} = \boldsymbol{\alpha}'\mathbf{z}_t + \boldsymbol{\beta}'\mathbf{x}_t + u_{t+h} \tag{1}$$

where y_{t+h} is the h-step-ahead value of the variable of interest (the real house price) which we want to forecast, u_{t+h} is a Gaussian forecast error distributed as $u_t \overset{iid}{\sim} N(0,\sigma^2)$, for $t = 1,...,T$, $\mathbf{z}_t$ is the $q \times 1$ vector of unrestricted predictors which are always included in the forecasting model, such as intercept and lags of the dependent variables, $\mathbf{x}_t$ is the $p \times 1$ vector of exogenous predictors whose dimensions we would like to shrink.

The unrestricted vector of coefficients $\boldsymbol{\alpha}$ and the variance σ^2 can be integrated out using the noninformative priors $\pi(\boldsymbol{\alpha}) \propto 1$ and $\pi(\sigma^2) \propto 1/\sigma^2$ respectively. This allows a closer focus on the regression coefficient vector $\boldsymbol{\beta}$, which has individual elements β_j, $j = 1,..., p.$.

A noninformative prior, like the one assigned to the coefficients $\boldsymbol{\alpha}$, leads to a Bayes estimator centered at the unrestricted least squares (LS) quantities. This choice obviously poses a problem for estimating the ‘‘large’’ number of coefficients $\boldsymbol{\beta}$, especially when $p > T$. This may lead to loss of degrees of freedom and poor forecast (curse of dimensionality problem). To solve this problem, some literature propose shrinkage of the predictors by computing the forecast as a projection on the first few principal components (De Mol et al., 2008) while others propose shrinkage using Bayesian approach (De Mol et al., 2008; Korobolis, 2013a, 2013b). The Bayesian approach uses either hierarchical or non-hierarchical priors. Starting with the hierarchical priors, the prior covariance matrix can be estimated in formal way, by placing hyper-prior distribution on its elements. Using the Bayes theorem, the hyper-prior times the likelihood will provide an appropriate posterior expression for the covariance matrix, which in turn can be used as the prior for the coefficients of interest $\boldsymbol{\beta}$.

As a generic example, the cases examined here can be cast into the prior form

$$\pi(\boldsymbol{\beta}|\tau^2) \sim N_p(0, V)$$

$$\pi(\tau_j^2) \sim F(\gamma),$$

where the $p \times p$ matrix V is the prior covariance matrix of the regression coefficients that we want to elicit for this ‘‘large p’’ problem. $F(\gamma)$ denotes a generic prior distribution on τ_j^2 with hyperparameter(s) γ. Following Korobilis (2013a), we also analyse five popular specifications of $F(\gamma)$ which are commonly used in the literature, leading to five Bayesian shrinkage estimators.

Hobert and Casella (1996) cited in Korobilis (2013a) studied the shrinkage properties of the Jeffreys prior on the covariance matrix of the regression coefficients. One can think of the Jeffreys prior as the simplest, default choice because it does not depend upon further hyperparameters. Let $V = \text{diag}\{\tau_j^2,...,\tau_p^2\}$; then the scale-invariant, improper Jeffreys hyperprior on each τ_j^2 takes the form

$$\pi(\tau_j^2) \sim 1/\tau_j^2, \text{ for } j = 1,..., p. \tag{2}$$

For a covariance matrix $V = \text{diag}\{\tau_j^2,...,\tau_p^2\};$, we can also consider a specific form of a gamma prior on τ_j^2, $j = 1,...p,$ i.e. the inverse gamma prior. This normal–inverse gamma mixture prior has been shown to be equivalent to a Student-t prior on $\boldsymbol{\beta}$ (Geweke, 1993 cited in Korobilis (2013a)). The t-density has heavy tails and is more leptokurtic around the origin, which means that shrinkage around zero is achieved at a faster rate than for the simple normal prior. The priors on τ_j^2 are of the form

$$\pi(\tau_j^2) \sim \text{igamma } (\rho, \xi), \text{ for } j = 1,..., p. \tag{3}$$

where ρ is the shape parameter and ξ the scale parameter of the inverse gamma density (Armagan and Zaretzki, 2010 cited in Korobilis (2013a))

Tibshirani (1996) cited in Korobilis (2013a) proposed the lasso algorithm, which can be viewed as a L_1-penalized least squares estimate. He noted that this form of penalized estimator is equivalent to the posterior mode of the Bayes estimate under the Laplace prior.

One can take advantage of the fact that the Laplace density can be written as a scaled mixture of normal (Park and Casella, 2008). This implies that for the Bayesian lasso prior (as well. as the fused lasso and the elastic net discussed below), we need to condition on the error variance σ^2. Park and Casella (2008) underline the fact that this conditioning ensures that the posterior of the regression coefficients $\boldsymbol{\beta}$ is unimodal, otherwise expensive simulation methods (for instance, simulated tempering) would be needed to handle multimodal posteriors. Assuming a diagonal prior covariance matrix of the form $V = \sigma^2 \times \text{diag}\{\tau_j^2,..., \tau_p^2\}$. The hierarchical version of the lasso augmented with the hyperprior is given as

$$\pi(\tau_j^2 | \lambda) \sim exponential\left(\frac{\lambda^2}{2}\right), \text{ for } j = 1,..., p. \tag{4}$$

where λ is a hyperparameter, which is the rate parameter of the exponential distribution i.e. the regularization parameter that controls the intensity of the penalty in each regression coefficient.

The hierarchical fused lasso which was proposed by Tibshirani et al. (2005) cited in Korobilis (2013a) as a means of accounting for any possible meaningful ordering of variables penalizes the L_1-norm of both the coefficients and their differences. Kyung et al. (2010) cited in Korobilis (2013a) show that the hierarchical representation of fused lasso prior is

$$\pi(\tau_j^2 | \lambda_1) \sim \text{exponential}\left(\frac{\lambda_1^2}{2}\right), \text{ for } j = 1,..., p. \tag{5a}$$

$$\pi(\omega_j^2 | \lambda_2) \sim \text{exponential}\left(\frac{\lambda_2^2}{2}\right), \text{ for } j = 1,..., p-1, \tag{5b}$$

In this case, V is a tridiagonal matrix, with main diagonal $\{\tau_i^2 + \omega_{i-1}^2 + \omega_i^2\}$ for $j = 1,..., p,$ and off-diagonal elements $\{-\omega_i^2\}$, and for simplicity we can set $\omega_0 = \omega_p = 0.$

Zou and Hastie (2005) cited in Korobilis (2013a) proposed the elastic net as a more stabilized version of the lasso that also allows grouping effects and is particularly useful when $P > T$. Kyung et al. (2010) cited in Korobilis (2013a)show that a hierarchical representation of the density exists, and that it is of double-exponential form, as in the simple lasso. This means that the hyperprior on τ_j^2 is

$$\pi(\tau_j^2 | \lambda_1) \sim \text{exponential}\left(\frac{\lambda_1^2}{2}\right), \text{ for } j = 1,..., p. \tag{6}$$

where, in this case, the difference from the standard lasso prior is that the covariance matrix is of the form $V = \sigma^2 \times \text{diag}\{(\tau_j^{-2} + \lambda_2)^{-1},..., (\tau_p^{-2} + \lambda_2)^{-1}\}$.

We use the Markov Chain Monte Carlo (MCMC) methods to obtain samples from the posterior distribution of all regression parameters $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2\}$. In particular, the Gibbs sampler is used to obtain draws from the posterior $p(\boldsymbol{\theta})$ by sampling from the conditional densities $p(\boldsymbol{\alpha} | \boldsymbol{\beta}, \sigma^2)$, $p(\boldsymbol{\beta} | \boldsymbol{\alpha}, \sigma^2)$ and $p(\sigma^2 | \boldsymbol{\alpha}, \boldsymbol{\beta})$.[1]

We also estimate bivariate predictive regressions between real house price and each of the 10 predictors. We include three lag of real house price as a control variable when testing the predictive ability of the specific predictor. We estimate the bivariate predictive regressions using ordinary least squares (OLS) and perform out-of-sample tests based on the recursive scheme.

Under the Gaussian prior, it is relatively simple to compute the maximiser of the posterior density, since, with independent and identically distributed (i.i.d.) regression coefficients, the solution amounts to solving a penalised least-squares of the coefficients (the Ridge regression problem). The double-exponential prior, on the other hand, does not have an analytical form for the maximiser of the posterior density, but under the prior of i.i.d. regression coefficients, the solution boils down to a Least Absolute Shrinkage and Selection Operator (Lasso) regression problem. Following De Mol et al. (2008) we also consider two algorithms for the Lasso regression, the least angle regression (LARS) and the iterative Landweber scheme with soft-thresholding at each iteration. Lasso regression combines variable selection and parameter estimation, with the estimator depending in a non-linear

[1] Details on specifications of the different minimization problems can be found in Korobilis (2013a). The MCMC procedure that we follow is also described in in the paper.

manner on the variable to be predicted.

We consider forecasts at one-step-ahead, of real house price growth rate using the relevant macroeconomic variables as predictors. Following standard practice, we use the model with no predictors (i.e. an autoregressive model) as a benchmark model. We evaluate the out-of-sample forecasts performance of the models using relative *MSE* which is the ratio of the unrestricted model's forecast mean squared error (*MSE*) to the restricted (AR) model's *MSE*. If the unrestricted model's *MSE* is less than the restricted model forecast *MSE*, this implies that the forecast produced by the former is better than the forecast from the later. Otherwise, if the unrestricted model's *MSE* is greater than the restricted model's forecast *MSE,* it implies that the simple *AR* model predicts more accurately than the unrestricted. To formally test whether forecasts from a specific model are significantly superior to the AR model forecasts, we use McCracken (2007) *MSE-F* statistic. The *MSE-F* statistic tests the null hypothesis that the unrestricted model forecast mean square error (*MSE*) is equal to the AR model forecast *MSE* against the one-sided (upper-tail) alternative that *MSE* of the specific model is less than the *MSE* of the AR model. Similar to the Diebold and Mariano (1995) statistic, the *MSE-F* statistic is based on the loss differential, and is given as:

$$MSE-F=(T-R-h+1).\bar{d}\,/\,MSE_1 \tag{7}$$

where T is the total sample, R is number of observations used for estimation of the model from which the first forecast is formed (i.e. the in-sample portion of the total number of observations), h is the forecast horizon , $M\hat{S}E_i=(T-R-h+1)^{-1}\sum_{t=R}^{T-h}(u_{i,t+1})^2$, $i=1,0$, $\bar{d}=M\hat{S}E_0-M\hat{S}E_1$, $M\hat{S}E_1$ corresponds to the *MSE* of the unrestricted model (i.e. model with the relevant macroeconomic variable (s) as predictors) and $M\hat{S}E_0$ corresponds to the *MSE* of the restricted model (i.e. the AR benchmark model). A significant *MSE-F* statistic indicates that the unrestricted model forecasts are statistically superior to those of the restricted model.

3. Data and Empirical Results

The dataset consists of 11 annual U.S. macroeconomic variables spanning the period 1890 to 2012. This include the real house price (*RHP*), the fiscal policy variable (*FISPOL*), real GDP per capita (*RGDPPC*), unemployment (*UNEMPL*), long term interest rate (*LTR*), short term interest rate (*STR*), inflation rate (*INFL*), population (*POP*), real construction cost (*RCONSTR*), real stock price (*RSP*)and real oil price (*ROILP*). All variables are from Robert J. Shiller web page, barring real GDP, population, unemployment and part of the budget surplus/deficit data used for computing the fiscal policy variable which are from the Global Financial data (GFD) base. For the budget surplus/deficit, we obtain the 1890 to 2006 data from GFD while 2007 to 2012 data are obtained from the Federal Reserve Bank, St. Louis. We use the ratio of budget surplus/deficit to GDP as our measure of fiscal policy. Inflation rate is computed as annual rate of growth in consumer price index. Real oil price is obtained by deflating the nominal West Texas Intermediate with CPI. Unemployment rate, fiscal policy, short and long term interest rates, and inflation rates were found to be stationary; hence they are expressed in levels. We use the first log differences (growth rates) for the remaining variables. All data series are plotted in Appendix 1.

We determine the optimal number of lags of the dependent variable to include in the forecasting model based on the Schwarz information criterion (SIC), which in turn, selected three lags. Hence, we always include the intercept and three lags of the dependent variables to capture the dynamics of real house price in the U.S.

The first estimation period (after taking lags and transforming to stationarity) is 1894 to 1916, given rise to 23 observations for the in-sample period. Our choice of this in-sample period is determined by test of multiple structural breaks for equation (1) using Bai and Perron (2003) approach. The test indicates two significant breaks occurred in 1917 and 1945.[2] Therefore, the sample 1917 to 2012 (last 96 observations) is kept for the evaluation of out-of-sample forecasts. All estimations are done recursively and all procedures are applied to standardized predictors for ease of variable selection and model comparison.

To tune the hyperparameters, we set penalization parameter selected at the beginning of the forecasting sample for the i.i.d. Gaussian prior (Ridge regression) to 0.46 (which was obtained from the OLS estimation of the predictive regression), to explain 86.77% of the in-sample variance (at the beginning of the forecasting exercise). Lasso with least angle regression, keeps the same number of selected predictors for each step of the out-of-sample, hence, we select 1 predictor at each step. Lasso with iterative Landweber, keeps the same penalization parameter for the each step of the out-of-sample, hence, we set the penalization parameter selected at the beginning of the forecasting sample to 1.28, to keep 1 predictor at the beginning of the forecasting exercise.

For the hierarchical priors we set the values of the relevant hyperameters and regularization parameters following Korobilis (2013a). Hierarchical priors have the advantage of allowing the data to determine the prior hyperparameter of interest (covariance matrix of the normal prior in our case). However, introducing a second layer of hierarchy (the gamma-type densities) means that at least one new hyperparameter is introduced barring the case of normal-uniform prior. For the normal–inverse gamma prior (Student-t), we use a more informative prior igamma $(\rho = 3, \xi = 0.001)$, which concentrates τ_j^2 around the neighbourhood of zero since for very low values, the inverse gamma becomes equivalent to a Jeffreys prior for

[2] The fact that we trimmed the sample 15 percent at both ends could have resulted in no break during the recent crisis. However, this is not a problem since any breaks in the out-of-sample period is taken care of by our recursive estimations.

τ_j^2. For the hierarchical lasso prior, a conjugate prior which would facilitate posterior computations is needed. Hence, we use the gamma prior on λ^2 (not λ), of the form $\pi(\lambda^2) \sim \text{gamma}(r, \delta)$. Similarly, an additional layer on the hyperparameters λ_1, λ_2 of the fused lasso and elastic net priors is of the form $\pi(\lambda_1^2) \sim \text{gamma}(r_1, \delta_1)$ and $\pi(\lambda_2^2) \sim \text{gamma}(r_2, \delta_2)$. Therefore, we set $r = \delta = 0.01$ (and similarly $r_1 = \delta_1 = r_2 = \delta_2 = 0.01$) as this will produce near-uniform (noninformative) priors on the hyperparameters λ, λ_1 and λ_2. We determine the optimal number of factors to include in the principal component regression using the scree plot which in this case yielded five factors.

We evaluate the forecasting performance of the models relative to an autoregressive order 3 (*AR*(3)) benchmark for a one-step-ahead forecasts using the root mean square error (MSE). Results are reported in Table 1. A number of observations emerge from these results. First, we observe that six out of the twenty forecasting models produced forecasts which are superior to the *AR*(3) benchmark. Interestingly, five of these models (*Gaussian*, *Lasso-Lars*, *Lasso-Landweber*, *Jefferys* and *Student-t*) belong to the Bayesian class while the sixth is the fiscal policy model estimated with ordinary least squares. It is also interesting to know that the bivariate least square models produced forecasts which are all better than the full least square model. This emphasizes the importance of shrinkage of predictors instead of using all predictors, some of which may actually not contribute to the forecast accuracy of the variable of interest. Overall, the *Lasso-Landweber* turn out to be the best out-of-sample forecasting model for the U.S. real house price with a 9.12 percent improvement in MSE of the *AR* (3) benchmark. This is followed by the *Gaussian* model with a reduction in the *AR*(3) model's MSE by 8.53 percent.

To see if these six models' forecasts are statistically superior to the *AR* (3) forecasts, we use the McCracken (2007) MSE-F statistics. Results show that these models MSE are significantly lower than the MSE of the *AR*(3) model. Given the overall performance of the *Lasso-Landweber* model, we also compare its forecast MSE relative to the other five good performing models. The *MSE-F* statistics is a one sided test and is also only appropriate for nested models. Given that the Lasso-Landweber does not nest the Gaussian, Lasso-Lars, Jefferys and Student-t models, we use the Diebold and Mariano (1995) test to check whether the Lasso-Landweber forecast is significantly better than forecasts from the other four models. The null hypothesis of equal forecast accuracy could not be rejected in any case. However, using the MSE-F statistic, given that Lasso-Landweber nests the *LS-Fispol* model, we find that the former predicts significantly better than the latter at 1 percent level of significance.

These findings suggest that it is more rewarding to include other predictors in forecasting real house price for the US economy than simply using only the previous values of the real house price. This leads us to the question of which exogenous predictors are most important drivers of the U.S. real house price? We observe that using the *Lasso-Landweber* model, only few variables drive the movement of the U.S. real house price over the entire sample. While the short term interest rate is the main driver over 1917 to 1938 periods, the real construction cost is over the 1957 to 1977 period. However, it is quite interesting to know that the fiscal policy variable is consistently the best predictor in the recent time (1978 to 2012).[3] This finding is consistent with the previous observation that barring the Bayesian models, only the *LS-Fispol* among the rest, performed better than *AR*(3) benchmark with a reduction in MSE of the later by 2.04 percent. This suggests the relative role of fiscal policy in predicting real house price compared to the lags of real house price in particular and the rest of the exogenous predictors in general.

Overall, these findings are intuitive given that theoretically, various channels exist whereby fiscal policy and monetary policies (short term interest rate) can affect the housing markets (Afonso and Sousa 2011, 2012). Fiscal policy affects housing markets directly through various taxes and subsidies on home ownership as well as indirectly through effects on macroeconomic variables that influence the housing market. For example, taxes on housing capital gains and the imputed rental housing value, fiscal subsidies and value added taxes (VAT) on purchases of new houses, and the tax deductibility of mortgage payments and housing rents can affect housing price via their effects on households' disposable income and the demand of houses. An indirect effect of fiscal spending through the wage bill and government infrastructure spending can lead to both increases and decreases in the demand for homes. More broadly, the deterioration of the fiscal stance and uncertainty about the long-run sustainability of public finances can affect long-term interest rates and negatively impinge on the financing conditions for mortgages, pushing housing price downwards.

Short term interest rate as a monetary policy variable, directly influences the user cost of housing capital, expectations of future house-price movements, and housing supply. It also indirectly influences the real economy through standard wealth effects from house price, balance sheet, credit-channel effects on consumer spending, and balance sheet, credit-channel effects on housing demand (Mishkin, 2007). For instance, interest rate reductions make it cheaper to finance housing. In theory, lower credit costs should stimulate the demand for housing, thus causing real estate prices to go up.

Therefore, the role of monetary and fiscal policy in explaining housing market developments cannot be neglected as this forecasting exercise has shown.

[3] Results on this are available from authors upon request.

Table 1. Evaluation for 1-step-ahead out-of-sample forecasts

MODEL	**Model Description**	**MSE**
LS_ALL	Least square estimation with all predictors	1.1228
LS-FISPOL	Least square estimation with fiscal policy	0.9796*
LS-INFL	Least square estimation with inflation rate	1.0089
LS-LTR	Least square estimation with long term interest rate	1.0034
LS-POP	Least square estimation with short term interest rate	1.0082
LS-RCONSTR	Least square estimation with real construction cost	1.0013
LS-RGDPPC	Least square estimation with real GDP per capita	1.0048
LS-ROILP	Least square estimation with real oil price	1.0102
LS-RSP	Least square estimation with real stock price	1.0029
LS-STR	Least square estimation with short term interest rate	1.0629
LS-UNEMPL	Least square estimation with unemployment rate	1.0099
PCR	Principal component regression	1.0494
GAUSSIAN	Bayesian estimate with Gaussian prior	0.9147*
LASSO-LARS	Bayesian estimate with lasso under least angle regression	0.9220*
LASSO-LANDWEBER	Bayesian estimate with lasso under iterative Landweber	**0.9088***
ADAPTIVE JEFFREYS	Bayesian estimate with adaptive Jeffreys prior	0.9309*
ADAPTIVE STUDENT-T	Bayesian estimate with adaptive student-t prior	0.9318*
HIRARCHICAL-LASSO	Bayesian estimate with hierarchical lasso prior	8.8735
HIRARCHICAL-FUSED LASSO	Bayesian estimate with hierarchical fussed lasso prior	9.7627
HIRARCHICAL-ELASTIC NET	Bayesian estimate with hierarchical elastic net prior	6.3109
LANDWEBER VS FISPOL	Comparison of lasso-landweber with fiscal policy model	0.9277*
LANDWEBER VS GAUSSIAN	Comparison of lasso-landweber with Gaussian prior	0.9936
LANDWEBER VS LARS	Comparison of lasso-landweber with lasso-LARS	0.9858

LANWEBER VS JEFFERYS	Comparison of lasso-landweber with Jefferys prior	0.9763
LANDWEBER VS STUDENT-T	Comparison of lasso-landweber with student-t prior	0.9754

Note: * indicates 1% level of significance for the MSE-F test.

Conclusion

This paper evaluates the ability of a number of predictive regression models in forecasting real house price growth in the United States (U.S.) based on an annual data set on 10 macroeconomic variables from 1890 to 2012. We consider 20 predictive regression models which are ordinary least square estimated using all 10 exogenous predictors, bivariate least square regressions with only one regressor at a time, principal component regression with five common factors, three non-hierarchical Bayesian shrinkage estimators (Gaussian prior, Lasso prior with least angle regression algorithm and lasso prior with iterative Landweber algorithm), five Bayesian hierarchical shrinkage priors (adaptive Jeffery, adaptive student-t, lasso, fused lasso and elastic net). Using both MSE and McCracken (2007) MSE-F statistics and based on an out-of-sample period of 1917 to 2012 for one-year-ahead forecast horizon, six of the models produced forecast that are superior to the *AR* (3) benchmark. These are the bivariate least square with the fiscal policy variable, the three non-hierarchical Bayesian models, and adaptive Jefferys and student-t models. Overall, the standard lasso prior estimated with the iterative Landweber algorithm is the best model for forecasting real house price for the U.S. economy with about 9 percent reduction in the MSE of the *AR*(3) benchmark model. In general, the non-hierarchical Bayesian shrinkage estimators perform better than their hierarchical counterparts. Also the results from Lasso estimated with Landweber algorithm clearly show the relative importance of the fiscal policy, monetary policy (short term interest rate) and real construction costs variables in predicting the U.S. real house prices. These findings suggest that in general, forecasting real house price in the U.S. is better with Bayesian models since these have the ability of handling uncertainties in both data and model specifications. Also including more information is in general more rewarding for forecasting real house price than relying on only previous lags of the real house price.

REFERENCES

[1] **Afonso, A. and Sousa, R.M. (2011),** ***What are the Effects of Fiscal Policy on Asset Prices*****.** *Economic Modelling*, 28, 1871–1890;

[2] **Afonso, A. and Sousa, R.M. (2012),** ***The Macroeconomic Effects of Fiscal Policy*****.** *Applied Economics*, 44, 4439–54;

[3] **Agnello, L. and Schuknecht, L. (2011),** ***Booms and Busts in Housing***

Markets: Determinants and implications. *Journal of Housing Economics*, 20, 171–190;
[4] **Bai, J. and Perron, P. (2003),** ***Computation and Analysis of Multiple Structural Change Models***. *Journal of Applied Econometrics*, 6, 72–78;
[5] **Beltratti, A. and C. Morana (2010),** ***International House Prices and Macroeconomic Fluctuations***. *Journal of Banking & Finance*, 34(3), 533–45;
[6] **Das, S., Gupta, R. and Kabundi, A. (2010),** ***The Blessing of Dimensionality in Forecasting Real House Price Growth in the Nine Census Divisions of the US.*** *Journal of Housing Research*, 19 (1), 89–109;
[7] **De Mol C., Giannone, D. and Reichlin, L. (2008),** ***Forecasting Using a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components?*** *Journal of Econometrics*, 146 (2), 318–328;
[8] **Diebold, F.X. and Mariano, R.S. (1995),** ***Comparing Predictive Accuracy*.** *Journal of Economics and Business Statistics*, 13, 253–263;
[9] **Dua, P., Nishita, R. and Sahoo, S. (2008),** ***Forecasting Interest Rates in India.*** *The Journal of Applied Economic Research*, 2(1), 1–41;
[10] **Gupta, R. (2013),** ***Forecasting House Prices for the Four Census Regions and the Aggregate US Economy in a Data-rich Environment.*** *Applied Economics*, 45:33, 4677-4697;
[11] **Gupta, R. and Das, S. (2010),** ***Predicting Downturns in the US Housing Market: A Bayesian Approach*.** *Journal of Real Estate Finance and Economics*, 41 (3), 294–319;
[12] **Gupta, R. and Kabundi, A. (2010), Forecasting Real US House Price: Principal Components versus Bayesian Regressions.** *International Business and Economics Research Journal*, 9 (7), 141–152;
[13] **Gupta, R and Kabundi, A. and Miller, S.M. (2011),** ***Forecasting the US Real House Price Index: Structural and Non-structural Models with and without Fundamentals***. *Economic Modelling*, 28 2013–2021;
[14] **Iacoviello, M. and Neri, S. (2010),** ***Housing Market Spillovers: Evidence from an Estimated DSGE Model*.** *American Economic Journal: Macroeconomics*, 2, 125–164;
[15] **Korobilis, B. (2013a),** ***Hierarchical Shrinkage Priors for Dynamic Regressions with Many Predictors.*** *International Journal of Forecasting*, 29, 43–59;
[16] **Korobilis, B. (2013b),** ***Bayesian Forecasting with Highly Correlated Predictors.*** *Economics Letters*, 118 148–150;
[17] **Leamer, E.E. (2007),** ***Housing is the Business Cycle.*** *NBER Working Paper No*. 13428. http://www.nber.org/papers/w13428;
[18] **McCracken, M.W. (2007**), ***Asymptotics for Out-of-sample Tests of Granger Causality***. *Journal of Econometrics*, 140, 719–752;
[19] **Mikhed, V. and Zemčík, P. (2009),** ***Do House Prices Reflect Fundamentals? Aggregate and Panel Data Evidence*.** *Journal of Housing Economics*, 18, 140–149;

[20] **Mishkin, F.S. (2007),** ***Housing and the Monetary Transmission Mechanism.*** *NBER Working Paper No*. 13518 October 2007;

[21] **Park, T. and Casella, G. (2008),** ***The Bayesian Lasso***. *Journal of the American Statistical Association*, 103, 681–686;

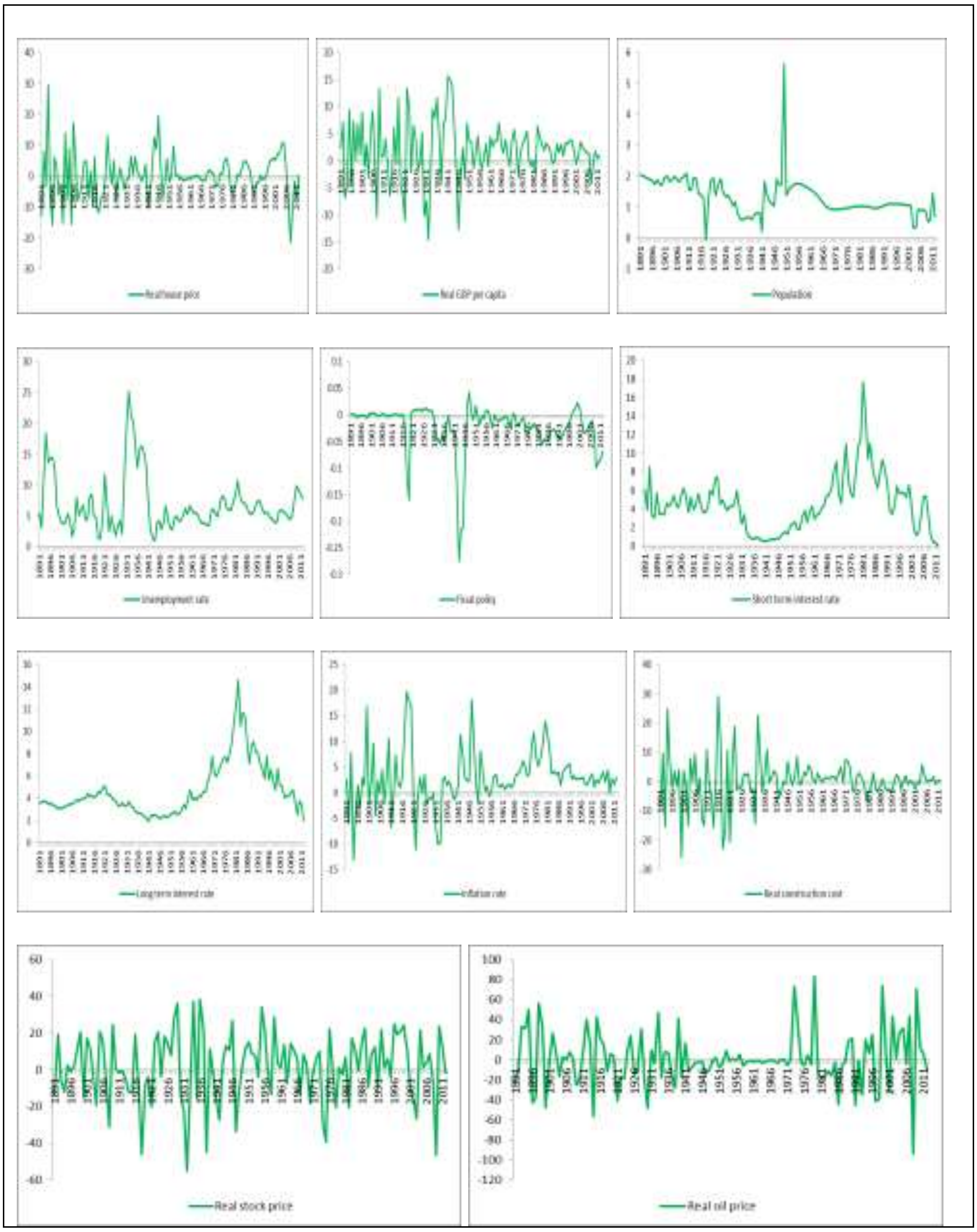
[22] **Rapach, D.E. and Strauss, J.K. (2007),** ***Forecasting Real Housing Price Growth in the Eighth District States*. *Federal Reserve Bank of St. Louis. Regional Economic Development***, 3(2), 33–42;

[23] **Rapach, D. E. Strauss, J. K. (2009),** ***Differences in Housing Price Forecastability across US States.*** *International Journal of Forecasting*, 25, 351–372;

[24] **Stock, J.H. and Watson, M.W. (2003),** ***Forecasting Output and Inflation: The Role of Asset Prices*.** *Journal of Economic Literature*, 41(3), 788-829;

[25] **Zeno and Füss (2010),** ***Macroeconomic Determinants of International Housing Markets***. *Journal of Housing Economics*, 19, 38–50.

Appendix 1. Data series



Note: All variables are in growth rates except unemployment rate, interest rates, inflation rate and fiscal policy.