Economic Modelling 61 (2017) 147–155

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/econmod

Forecasting house prices using dynamic model averaging approach: Evidence from China

Yu Wei[*], Yang Cao

School of Economics & Management, Southwest Jiaotong University, China

ARTICLE INFO

JEL:
C53
C22
G14
G13
R21
R31

Keywords:
House price forecasting
DMA
Rolling forecasting
MCS

ABSTRACT

Forecasting house price has been of great interests for macroeconomists, policy makers and investors in recent years. To improve the forecasting accuracy, this paper introduces a dynamic model averaging (DMA) method to forecast the growth rate of house prices in 30 major Chinese cities. The advantage of DMA is that this method allows both the sets of predictors (forecasting models) as well as their coefficients to change over time. Both recursive and rolling forecasting modes are applied to compare the performance of DMA with other traditional forecasting models. Furthermore, a model confidence set (MCS) test is used to statistically evaluate the forecasting efficiency of different models. The empirical results reveal that DMA generally outperforms other models, such as Bayesian model averaging (BMA), information-theoretic model averaging (ITMA) and equal-weighted averaging (EW), in both recursive and rolling forecasting modes. In addition, in recent years it is found that the Google search index, instead of fundamental macroeconomic or monetary indicators, has developed greater predictive power for house price in China.

1. Introduction

House prices are important indicators of the real estate market's health and stability. Forecasting the changes in house prices can help in gaining a clear understanding of the real estate market. For this reason, house investors, real estate developers and government regulators always pay great attention to the trends in house prices. Forecasting house prices accurately can not only help the government to regulate the real estate market more effectively, but it can also help real estate developers to make their investment decisions properly. House investors' decisions also depend largely on predictions of future house prices. The more accurate the predictions of future house prices, the more rational the buyers can be in allocating their current and future consumptions.

A number of scholars have conducted extensive research in predicting house prices. For example, DiPasquale and Wheaton (1994) explore the dynamic mechanism of house prices in the United States during the 1980s. These authors apply several macroeconomic variables for the first time to forecast house prices, and they find that these macroeconomic variables can improve the accuracy of house price forecasts. Brown et al. (1997) use the time-varying parameter (TVP) model to forecast quarterly changes in house prices in the U.K. from 1968 to 1992. They show that the TVP model performs better than several traditional constant parameter regression models, e.g., the error correction mechanism, the vector autoregressive and the autoregressive time series regression models. Crawford and Fratantoni (2003) forecast quarterly house prices in five U.S. states from 1979 to 2001 by using the ARIMA, GARCH and transition matrix methods. Their study reveals that according to the in-sample goodness-of-fit test, the transition matrix model performs better than other models, as it allows the variable parameters to change over time. However, for out-of-sample forecasting, the traditional ARIMA model is superior to the others. Hadavandi et al. (2011) investigate the annual changes in house prices in 20 regions of Iran since 2001 by using a fixed effects model with panel data. They find that the forecast accuracy of the panel data model is better than that of the ordinary OLS regression. Rapach and Strauss (2009) study the quarterly house prices of the top 20 major cities in the U.S. from 1995 to 2006, and then compare the differences in predictions produced by auto-regression (AR) models and ARDL models that incorporate information from numerous economic variables. Rapach and Strauss find that the AR models often provide relatively accurate house price forecasts in a number of interior states. However, all of the forecasting models tend to perform poorly for a group of primarily coastal states that have experienced especially strong growth in house prices. Ghysels et al. (2013) classify the recent literature on house price forecasting, and summarize the defects and problems of traditional forecasting models.

However, those traditional house price forecasting models face at

[*] Correspondence to: First Section of Northern Second Ring Road 111, Chengdu, China.
E-mail address: weiyusy@126.com (Y. Wei).

http://dx.doi.org/10.1016/j.econmod.2016.12.002
Received 10 August 2016; Received in revised form 29 November 2016; Accepted 2 December 2016

least one or two main problems as follows: first, it has been documented in the recent literature that the effects of the determinants (predictors) on house prices change over time (Ghysels et al., 2013). Lots of factors such as macroeconomic cycles or policy adjustments on real estate industry may lead to the structural break in the relationship between fundamentals and house price dynamics. Furthermore, the influence of each determinant on house prices may not be identical during different periods of time and (or) under different market conditions (Rapach and Strauss, 2009; Ghysels et al., 2013; Nneji et al., 2013; Plakandaras et al., 2015). Second, it is found that one specific model with a fixed set of predictors might not perform well consistently over time. To handle this issue, a model selection procedure may be carried out at each point in time, but with a huge computational task. For example, if there are n predictors in hand, that is, 2^n models are needed to be assessed at each point in time and a total number of models to be assessed over the evaluation period of T is 2^{nT}. If both n and T are large, this task seems unachievable. Thus the model averaging method such as forecast combination is adopted to improve the forecasting accuracy. When considering different sets of predictors as separate models, model averaging is simply a weighted average of all possible combinations of predictors. However, the weights assigned by either simple forecast combination or Bayesian model averaging (BMA) to combine different models are constant over time, which is not flexible enough to capture the time-variation of the contribution from each model (Próchniak and Witkowski, 2013; Man, 2015). Recently, Kapetanios et al. (2008) propose a new model combination method, named as information-theoretic model averaging (ITMA). This method adopts Akaike information criterion (AIC) of each individual model in previous observations to update the model probability by the ordinary BMA. The ITMA method is found to be a powerful alternative to Bayesian averaging schemes (Piegorsch et al., 2013).

To address the problems mentioned above, Raftery et al. (2010) propose a novel method known as dynamic model averaging (DMA). It is found that this approach works quite well in macroeconomic forecasting (see, e.g., Koop and Korobilis, 2011, 2012). There are several obvious advantages of the DMA approach. It allows both the sets of predictors (forecasting models) and the coefficients of predictors to change over time. Furthermore, DMA method combines models in a dynamic way using two forgetting factors to approximate the evolution of model parameters and model switching probabilities, respectively. These two forgetting factors make the heavy computational task of model selection procedure manageable.

Motivated by the work of Raftery et al. (2010) and Koop and Korobilis (2012), a study by Bork and Møller (2015) is the first application using the dynamic model averaging (DMA) and dynamic model selection (DMS) methods in forecasting house price. Their study investigates the predictability of quarterly house price data in 50 U.S. states from 1976 to 2012. Unlike the traditional constant parameter or TVP forecasting models, the DMA and DMS methods allow changes over time in both the forecasting models and their coefficients. Bork and Møller (2015) also find that the forecasting precision of the DMA model is about 30% better than that of the traditional time series methods such as the AR and OLS regression models. More recently, Akinsomi et al. (2016) use DMA to forecast the growth rates of U.S. real estate investment trusts (REITs) from January 1991 to December 2014. They find that compared to the traditional predictors, the sentiment and uncertainty indicators can better predict changes in REITs. DMA can also outperform other models such as Bayesian model averaging (BMA) and AR, as DMA produces smaller forecasting errors. Their results suggest that economy-wide indicators, monetary policy instruments and sentiment indicators are among the most powerful predictors of REIT returns. However, it is very interesting to note that besides those fundamental macroeconomic and monetary policy indicators, with the recent dramatic development of Internet information and big-data technology, there is growing evidence that Internet online search indexes can be useful in predicting some future economic variables, social population features and even outbreaks of epidemic diseases. However, to the best of our knowledge, there is no direct research focusing on the relationship between Internet search indexes and changes in house prices. A relevant paper by Das et al. (2015) examines the association between online apartment rental searches and several fundamental real estate market variables. These researchers find that consumers' online Google search indexes are significantly associated with vacancy rates, rental rates and U.S. REIT returns.

It is also important to note that in most of the previous house price forecasting studies, different models are evaluated simply by a single loss function, e.g., the mean squared forecast error (MSFE) or mean absolute forecast error (MAFE). However, when a particular loss function value is smaller for model A than it is for model B, we cannot arbitrarily conclude that the forecasting performance of model A is superior to that of model B. Such a conclusion cannot be made on the basis of a single loss function or a single data sample.

Recent work has focused on a testing framework that can determine whether one particular model outperforms another (Diebold and Mariano, 1995; West, 1996; White, 2000). Hansen et al. (2011) propose a new statistical test for forecast errors, namely the model confidence set (MCS) test. The MCS test has several attractive advantages over conventional tests such as the super predictive ability test of Hansen (2005) or the reality check test by White (2000). First, the MCS test does not require a benchmark model to be specified, which is very useful in applications without an obvious benchmark. Second, the MCS test acknowledges the influence of measurement or calculation errors in a data sample by using a bootstrap sampling technique. Third, the MCS procedure allows for the possibility of more than one "best" model.

Since 2007, the real estate market in large and medium-sized Chinese cities has entered a stage of high-speed growth in which house prices have risen sharply. For example, the national average of house prices has increased by 7.6% per year, and some developed cities have even seen sustained double-digit growth in prices. The overheated real estate industry with its high house prices not only traps the economic capital of property-owning entities, but it also brings the disadvantages of excessive dependence on land finance for government income and other economic or social problems. In 2013, the government began to introduce control measures to prevent house prices from rising too quickly, but these measures appear to have little effect. The "China City Life Quality Report 2013" issued by the Experimental Research Institute of Chinese Economy, indicated that more than 40% of China's urban residents felt the city house prices were unacceptably high. Thus accurate forecasting of house price in China is of great importance for policy makers and different market participants.

As discussed above, this paper contributes to the literature in four ways. First, we forecast house price in major Chinese cities using dynamic model averaging (DMA) for the first time, and compare its performance with that of Bayesian model averaging (BMA), information-theoretic model averaging (ITMA) and other commonly-used forecasting models. Second, unlike the method used by Bork and Møller (2015) and Akinsomi et al. (2016), this investigation applies both recursive forecasting and rolling-window prediction in the DMA forecast to enhance the robustness of empirical performance for different models. Although the recursive method is usually used in DMA forecasting and the forgetting factor approach is able to play a role in discarding early observations and giving more weights to the most recent ones, it cannot handle the problem of possible structural break in house price dynamics. Third, this paper is the first to use a Google online search index as an additional predictor, along with several fundamental macroeconomic and real estate market indicators, to forecast house price changes in major Chinese cities. Fourth, instead of depending on one single loss function criterion, this paper applies a newly developed model comparison technique, the MCS test, which is proposed by Hansen et al. (2011) to assess the forecasting performance of different models using a number of criteria and test statistics.

The structure of the rest of the paper is as follows. Section 2 provides the methodology of DMA forecasting. Section 3 describes our data sample and lists the summary statistics of the data. Section 4 provides the empirical results for different forecasting models, and Section 5 presents the conclusions.

2. Methodology of DMA (DMS) and model evaluation

2.1. DMA and DMS

Most traditional time-series forecasting models, such as multi-variable regression or AR (auto-regression), are constant coefficient (CC) models. Despite their advantages of providing simple and easy estimation, their drawbacks are also obvious. To be specific, the regressor coefficients are fixed, and are not allowed to change over time. Hence, the time-varying parameter (TVP) model has emerged to overcome the flaws of CC models. The TVP method allows the parameters of explanatory variables to change over time, incorporating the naturally time-varying relationship between dependent and independent variables. As noted by Primiceri (2005) and Koop et al. (2009), the ordinary TVP model can be presented as follows:

$$y_t = x'_{t-1}\beta_t + \varepsilon_t, \tag{1}$$

$$\beta_t = \beta_{t-1} + \eta_t, \tag{2}$$

where y_t is the dependent variable to be forecasted. In our paper, y_t is the log monthly rate of house price growth in each selected major city in China at time t. x_{t-1} is a $1\times m$ vector of predictors. β_t is an $m\times 1$ vector of coefficients, and the innovation items are distributed as $\varepsilon_t \sim i.\ i.\ d.\ N(0, V_t)$, $\eta_t \sim i.\ i.\ d.\ N(0, W_t)$. This kind of TVP model can be estimated by the Kalman filter method.

In the TVP model as defined by Eqs. (1) and (2), it is assumed that the predictors in x_{t-1} are fixed throughout all of the time points, which may lead to a substantial loss of forecasting precision and problems of over-parameterization. However, DMA and DMS improve the TVP model by allowing the predictor sets (forecasting models) and their coefficients to both change over time. Therefore, following the novel work of Raftery et al. (2010) and Koop and Korobilis (2012), this paper uses DMA and DMS methodologies to forecast house price growth in major Chinese cities. The DMA (DMS) method is illustrated as follows:

$$y_t = x_{t-1}^{(k)\prime}\beta_t^{(k)} + \varepsilon_t^{(k)}, \tag{3}$$

$$\beta_t^{(k)} = \beta_{t-1}^{(k)} + \eta_t^{(k)}, \tag{4}$$

where $x_{t-1}^{(k)\prime} \subseteq x'_{t-1}$ for k=1, 2, ..., K denotes a specific predictor set, and $\varepsilon_t \sim i.\ i.\ d.\ N(0, V_t^{(k)})$, $\eta_t \sim i.\ i.\ d.\ N(0, W_t^{(k)})$. If there are m predictors in x'_{t-1}, the total number of possible combinations of these predictors (possible forecasting models) will be K=2^m. DMA and DMS can then incorporate the uncertainty factors from these K models' in a dynamic way:

$$\hat{y}_t^{DMA} = \sum_{k=1}^{K} \pi_{(t|t-1,k)} x_{t-1}^{(k)\prime}\beta_{t-1}^{(k)}, \tag{5}$$

$$\hat{y}_t^{DMS} = x_{t-1}^{(k^*)\prime}\beta_{t-1}^{(k^*)}, \tag{6}$$

where the probability of model k is $\pi_{(t|t-1,k)} = \Pr(L_t = k|Y^{t-1})$. The equation L_t=K indicates that model K is selected at time t, and $Y^{t-1} = \{y_1, \ldots, y_{t-1}\}$. DMA obtains the forecasting result at any point in time by taking the average of all the K models according to their historical forecasting performances, denoted by $\pi_{(t|t-1,k)}$. DMS then chooses the model with the best historical performance, i.e., with the highest probability $\pi_{(t|t-1,k^*)}$.

Raftery et al. (2010) propose a simple estimation of DMA. This estimation simplifies calculation without loss of forecasting accuracy by using the Kalman filter method. The initial assumptions of this estimation are that $\beta_{t-1}^{(k)}$ is independent and identically distributed, and that $\beta_{t-1}^{(k)}$ can be determined separately only if $L_{t-1} = k$. In this set of assumptions, λ is recommended as a so-called forgetting factor. λ plays the most important role in the calculation of $\beta_{t-1}^{(k)}$, where j period gains λ^j weight from the starting period. λ can also simplify the covariance matrix of $\beta_{t-1}^{(k)}$. This process is given below:

$$\beta_{t|t-1}^{(k)} = \beta_{t-1|t-1}^{(k)}, \tag{7}$$

$$\Sigma_{t|t-1}^{(k)} = \frac{1}{\lambda}\Sigma_{t-1|t-1}^{(k)}, \tag{8}$$

where $\Sigma_{t|t-1}^{(k)}$ denotes the covariance matrix of $\beta_{t-1}^{(k)}$. Then, the estimation of the DMA parameters is achieved by the following updating equations:

$$\hat{\beta}_{t|t}^{(k)} = \hat{\beta}_{t-1|t-1}^{(k)} + \Sigma_{t|t-1}^{(k)} x_{t-1}^{(k)\prime}\left(V_t^{(k)} + x_{t-1}^{(k)\prime}\Sigma_{t|t-1}^{(k)} x_{t-1}^{(k)}\right)^{-1}\left(y_t - x_{t-1}^{(k)\prime}\hat{\beta}_{t-1}^{(k)}\right), \tag{9}$$

$$\Sigma_{t|t}^{(k)} = \Sigma_{t|t-1}^{(k)} - \Sigma_{t|t-1}^{(k)} x_{t-1}^{(k)\prime}\left(V_t^{(k)} + x_{t-1}^{(k)\prime}\Sigma_{t|t-1}^{(k)} x_{t-1}^{(k)}\right)^{-1} x_{t-1}^{(k)}\Sigma_{t|t-1}^{(k)}, \tag{10}$$

where Eq. (8) is simplified with $W_t^{(k)} = (\lambda^{-1} - 1)\Sigma_{t|t-1}^{(k)}$ by using the forgetting factor.

Another forgetting factor is used in the second assumption, in which α is used in Eqs. (5) and (6). If we use a transition matrix of probability, we must consider $K = 2^m$ model combinations with m predictors at each time point. Once m is larger than 5, it is not practicable to operate the Markov switching in the $K\times K$ matrix. Thus, using the forgetting factor α is a practical way to reduce calculation time and error. In this way, the probability in the forecasting model is determined as follows:

$$\pi_{(t|t-1,k)} = \frac{\pi_{(t-1|t-1,k)}^{\alpha}}{\sum_{\ell=1}^{K}\pi_{(t-1|t-1,\ell)}^{\alpha}}, \tag{11}$$

and the updating equation is

$$\pi_{(t|t,k)} = \frac{\pi_{(t|t-1,k)} f_k(y_t|Y^{t-1})}{\sum_{\ell=1}^{K}\pi_{(t|t-1,k)} f_\ell(y_t|Y^{t-1})}, \tag{12}$$

where $f_\ell(y_t|Y^{t-1})$ denotes the predictive density of model ℓ. The Eqs. (7)–(12) consist of all the steps of the Kalman filter prediction and the updating process. Raftery et al. (2010) indicates that if λ=α=1, then DMA can be treated as BMA (Bayesian model averaging) without any forgetting.

Furthermore, Raftery et al. (2010) and Koop and Korobilis (2012) use a constant forgetting factor of λ=α=0.99 in each period. However, in Eqs. (7)–(12), the value of a forgetting factor determines how rapidly the forecasting models and parameters evolve. We can see that it is suitable to permit the models' probabilities to change more quickly in some time periods than in others, and to permit the models' parameters to adjust according to their longer historical performance more than according to their most recent behavior. However, if we set the same constant forgetting factors at all time points, there is a major loss of predictive accuracy. Thus, it is necessary to allow DMA to choose its own proper forgetting factors to minimize the forecasting errors at each time point. Following Bork and Møller (2015), we first apply the dynamic forgetting factor α_t as

$$\arg\max_{\alpha_t} \sum_{k=1}^{K} f_k(y_t|Y^{t-1}) \frac{\pi_{(t-1|t-1,k)}^{\alpha_t}}{\sum_{\ell=1}^{K}\pi_{(t-1|t-1,\ell)}^{\alpha_t}}, \tag{13}$$

where α_t takes a finite number of values. We run a Kalman filter progress for each α_t value. Then, a value of α_t is selected that produces the K models with the highest probability in each period. As it is adopted by Koop and Korobilis (2012), α_t can only take the five values in the interval of $\alpha_t \in \{0.95, \ldots, 0.99\}$. Second, the typical value of λ is determined by the most recent forecasting errors (Fortescue et al., 1981). However, we set a time-varying $\lambda_t^{(k)}$ by taking into account the performance of the model k during the whole forecasting period. This time-varying λ_t can reduce the noise of unexpected big forecasting errors in some periods. We first calculate Ψ quantiles of squared forecasting errors as $\varepsilon_k^{t-1} = \{\varepsilon_{k,1}^2, \dots, \varepsilon_{k,t-1}^2\}$ across the whole forecasting period. Then, we denote its quantiles as $q_\psi(\varepsilon_k^{t-1})$, $\psi = 1, \dots, \Psi$. Next, if $\varepsilon_{k,t}^2$ is in the same interval of $[q_{\psi-1}(\varepsilon_k^{t-1}), q_\psi(\varepsilon_k^{t-1})]$ as in the previous period, then $\lambda_t^{(k)} = \lambda_{t-1}^{(k)}$. Otherwise, $\lambda_t^{(k)}$ increases or declines with the current forecasting errors. Finally, in accordance with Koop and Korobilis (2012), we limit the values of $\lambda_t^{(k)}$ with the same interval of α_t to reduce the computational burden.

2.2. Model evaluation

To evaluate the forecasting performances of different models, various criteria may be used. To get more robust conclusions on the accuracy of these models, we use two loss functions as our evaluation criteria. These loss functions are the MSFE (mean squared forecast error) and the MAFE (mean absolute forecast error), which are defined as follows:

$$\mathrm{MSFE} = M^{-1}\sum_{t=1}^{M}(y_t - \hat{y}_t)^2, \tag{14}$$

$$\mathrm{MAFE} = M^{-1}\sum_{t=1}^{M}|y_t - \hat{y}_t|, \tag{15}$$

where y_t is the true house price growth rate, and $\hat{y}_t$ is the set of forecasts for the house price growth rate made by different forecasting models. M is the number of forecasts.

However, the abovementioned loss functions do not provide any information on whether the differences of forecasting losses among models are statistically significant. Therefore, to choose the superior models, we use an advanced statistical test, namely the MCS test as proposed by Hansen et al. (2011). The MCS method has several attractive advantages over conventional tests such as those recommended by Diebold and Mariano (1995), West (1996), White (2000) or Hansen and Lunde (2005). First, the MCS test does not require a benchmark model to be specified, which is very useful in applications without an obvious benchmark. Second, the MCS test acknowledges the likelihood of outliers in the data, which may cause extreme bias of loss functions. Third, the MCS test allows for the possibility of more than one "best" model. The MCS test is processed as follows.

Consider a set, M_0, that contains a finite number of objects, indexed by m=1, ..., m_0. The objects are evaluated over the sample t=1, ..., n in terms of a loss function i, and we denote the loss that is associated with object u in period t as $L_{u,t}$. We define the relative performance variables $d_{uv,t} \equiv L_{u,t} - L_{v,t}$ for all u, $v \in M_0$. The set of superior objects is then defined by

$$M^* \equiv \{u \in M_0 : E(d_{i,uv,t}) \le 0 \quad for\ \ all\ \ v \in M_0\}, \tag{16}$$

where $E(d_{i,uv,t})$ is the mathematical expectation of $d_{uv,t}$ under a specific loss function i, such as MSFE or MAFE. The MCS test is done through a sequence of significance tests in which objects that are found significantly inferior to other elements of M_0 are eliminated. The null hypotheses that are being tested take the form of

$$H_{0,M}: E(d_{i.uv,t}) = 0 \qquad for\ \ all\ \ u, v \in M \subset M_0. \tag{17}$$

The MCS procedure is based on an *equivalence test*, δ_M, and an *elimination rule*, e_M. The equivalence test, δ_M, is used to test the hypothesis H_0 for any $M \subset M_O$, and e_M identifies the object of M that is to be removed from M in the event that H_0 is rejected. After repeating these two tests, we obtain the set, $\widehat{M}_{1-\alpha}^*$, which consists of the set of "surviving" objects (those that have survived all tests without being eliminated), and which are referred to as the MCS (model confidence set). In Hansen et al. (2011), the significance level α is set to 0.1. This significance level means that if the p-value is larger than 0.1, then the corresponding model is a "surviving" model, implying that its forecasting performance is relatively good.

The range statistic (T_R) and the semi-quadratic statistic (T_{SQ}), which are suggested by Hansen et al. (2011), are commonly used in model evaluation. These variables are defined as follows:

$$T_R = \max_{u,v\in M}\frac{|\bar{d}_{i,uv}|}{\sqrt{\mathrm{var}(d_{i,uv})}}, \quad T_{SQ} = \max_{u,v\in M}\frac{(\bar{d}_{i,uv})^2}{\mathrm{var}(\bar{d}_{i,uv})}, \tag{18}$$

where $\bar{d}_{i,uv} = \frac{1}{n}\sum_{t=1}^{n} d_{i,uv,t}$. If the p-value of T_R and $T_{SQ}$is larger than 0.1, we conclude that the null hypothesis (Eq. (17)) cannot be rejected. The asymptotic distributions of the test statistics T_R and T_{SQ} are non-standard, because they depend on nuisance parameters (under both the null and the alternative conditions). However, these conditions pose no obstacle, as the distributions of these statistics are easily estimated by using bootstrap methods that implicitly solve the nuisance parameter problem. In addition to T_R and T_{SQ}, we use four more test statistics, i.e., T_{max}, T_Q, T_F and T_D to obtain more robust conclusions for model comparison. More details on these test statistics are discussed in Hansen et al. (2011).

3. Data and descriptive statistics

We collect the monthly new commodity house price indices of 30 provincial capitals and municipalities in China between January 2007 and December 2015. These data are available from the website of the National Bureau of Statistics of China (http://data.stats.gov.cn/index.htm). The growth rates of house prices for each of these 30 cities are calculated as

$$y_{i,t} = 100 \times [\ln(price_{i,t}) - \ln(price_{i,t-1})]\, i = 1, \dots, 30, \tag{19}$$

where $price_{i,t}$ denotes the new commodity house price indices, adjusted by CPI for city i at time t. Furthermore, following Hamilton and Schwab (1985), Case and Shiller (1990), Malpezzi (1999), Rapach and Strauss (2009) and Bork and Møller (2015), the fundamental economic and monetary predictors for house price growth rates are chosen for two categories: the national level and city level.

As discussed above, there is growing evidence that Internet online search indexes can be useful in predicting some future economic variables. Thus, in addition to the economic and social predictors commonly used in house price forecasting, we adopt a new predictor – the Google online search index. This index is based on searching for the items "city name + house price" in China. All of the variables involved are listed and explained in Table 1.

In our analysis, the annual or quarterly data, i.e., the log growth

Table 1
Dependent and independent variables of the forecasting model.

Variable names	Definitions (percentage)	Level	Frequency
y	log growth rate of real house price	City	Monthly
cpi	log growth rate of consumer price index	City	Monthly
gdp	log growth rate of gross domestic product	City	Monthly
hinvest	log growth rate of investment in real estate market	City	Monthly
income	log growth rate of real disposable income per capita	City	Quarterly
unemploy	log growth rate of unemployment	City	Annually
search	log growth rate of Google online search index	City	Monthly
housestarts	log growth rate of newly built house areas	Nation	Monthly
consum	log growth rate of total retail sales of consumer goods	Nation	Monthly
indusgr	log growth rate of industrial production growth	Nation	Monthly
interspr	the spread between 10-year and 3-month Treasury rates	Nation	Monthly
loanr	the 30-year mortgage rate	Nation	Monthly
prosper	log growth rate of the national real estate prosperity index	Nation	Monthly

Table 2
Descriptive statistics of dependent and independent variables for Beijing.

Variables	Obs.	Mean	Min.	Max.	Std. dev.	ADF
y	108	0.5904	−1.3085	4.4017	0.9112	−3.1322**
cpi	108	0.0007	−2.0359	1.4444	0.6382	−3.4312**
gdp	108	1.3288	−8.7314	9.1207	5.0321	−4.2590***
hinvest	108	0.0498	−1.7430	3.4207	0.6672	−7.3825***
income	108	0.7671	−3.7218	5.9096	2.6295	−3.9609***
unemploy	108	−0.3255	−1.9650	0.4204	0.6564	−3.0503*
search	108	−0.0367	−0.8109	0.5679	0.2676	−10.9497***
housestarts	108	0.0001	−1.8508	1.2242	0.6445	−9.7524***
consum	108	0.1162	−17.9901	14.2284	7.6478	−12.6419***
indusgr	108	−0.6819	−58.2169	65.7429	15.2903	−11.2713***
interspr	108	−1.0833	−97	98	27.3466	−9.4779***
loanr	108	0.0045	−16.2519	6.8053	2.2852	−7.0455***
prosper	108	−0.0636	−2.052	2.4083	0.7472	−4.2769***

Note: ADF means the statistics from the augmented Dickey-Fuller unit root tests. The asterisks *, ** and *** denote rejections of null hypothesis at the 10%, 5% and 1% significance levels, respectively.

rate of unemployment and log growth rate of real disposable income per capita, are converted into monthly data by a linear interpolation algorithm. Taking Beijing as an example, we present the descriptive statistics of all these variables in Table 2.

Beijing, the capital of China, is very famous for its high levels and rates of growth in house prices. Table 2 shows that the average monthly growth rate for house prices in Beijing is about 0.59%, as compared to the average growth rate for all of the selected 30 major cities in China, which is about 0.36%. During the sample period, the highest average house price growth rate is recorded at about 0.66% in the city of Haikou, the capital of Hainan province. The lowest rate is observed at about 0.18% in the city of Hohhot, the capital of Inner Mongolia autonomous region. Table 2 also shows that the volatility for these predictors is quite different for Beijing. For example, the standard deviation for the growth rate in total retail sales of consumer goods is about 7.65%, whereas this figure for the growth rate on the Google online search index is just 0.27%. Furthermore the ADF unit root tests suggest that all of the time series are stationary, indicating that we may model them through econometrics analysis without further transformation.

4. Empirical analysis

4.1. Prediction accuracy test

In the extant literature on house price forecasting, recursive forecasting techniques are commonly applied in which all of the historical data before the forecasting time point are assumed to form a new forecast (for example, see Bork and Møller, 2015 and Akinsomi et al., 2016). However, in most time series forecasting studies, rolling-window forecasting is recognized as a more accurate methodology than recursive methods, due to the possibility of structural breaks in time series. Thus, we use both recursive and rolling time-window forecasts of house prices to draw the most robust conclusions on performance. Within the recursive forecasting mode, we can obtain the forecasting results for almost the whole data sample period, i.e., March 2007 to December 2015. However, with the rolling-window forecasting mode, we fix the in-sample length as six years, and the out-of-sample data are chosen from the last three years of the whole sample (or for 2013–2015). In other words, we investigate two different forecasting modes and evaluation samples, which implies that our empirical results are quite robust and reliable.

In the following tests, the one-month ahead house price growth rate is predicted by different models using recursive and rolling-window forecasting modes. The models used in this horse race are as follows:

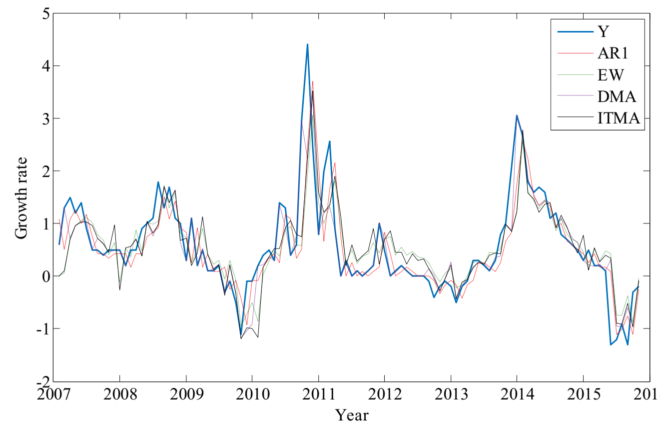

Fig. 1. Recursive forecasting results from different models for Beijing.

1) EW: equal-weighted averaging of K OLS models, i.e., the equal-weighted DMA model with recursive mode;
2) AR1: first-order autoregression model on house price growth rate with recursive mode;
3) DMA: dynamic model averaging with recursive mode;
4) DMS: dynamic model selection with recursive mode;
5) BMA: DMA prediction with recursive mode, and $\lambda = \alpha = 1$;
6) BMS: DMS prediction with recursive mode, and $\lambda = \alpha = 1$;
7) ITMA: information-theoretic model averaging with recursive mode;
8) EW_ROLL: EW with rolling-window forecasting mode. The rolling window is fixed as six years, and the years 2013 to 2015 are taken as the out-of-sample period;
9) AR1_ROLL: AR1 with rolling-window forecasting mode;
10) DMA_ROLL: DMA with rolling-window forecasting mode;
11) DMS_ROLL: DMS with rolling-window forecasting mode;
12) BMA_ROLL: BMA with rolling-window forecasting mode;
13) BMS_ROLL: BMS with rolling-window forecasting mode;
14) ITMA_ROLL: ITMA with rolling-window forecasting mode.

To be clear, Fig. 1 presents the recursive forecasting results of only four models for Beijing. It is shown that the real house price growth rate for Beijing (Y) is quite volatile throughout the 2007–2015 period. In general, however, all of the four forecasting models follow the trends of house price changes in Beijing with small discrepancies. Thus, to quantitatively compare the forecasting accuracy of different models, it is preferable to use more statistical methods.

To save space, Table 3 reports only the mean square forecast error

Table 3
Summary for MSFE for all 30 selected major cities in China.

Model	Mean	Std. dev.	Min.	Max.	Obs.
Panel A: Recursive mode					
EW	0.7788	0.9098	0.1950	5.0063	106
AR1	0.5222	0.6079	0.1198	3.4995	106
DMA	0.2759	0.2258	0.0529	1.0528	106
DMS	0.3410	0.2621	0.0890	1.2842	106
BMA	0.6112	0.8638	0.1281	4.9402	106
BMS	0.2833	0.2200	0.0576	1.0084	106
ITMA	0.3365	0.2099	0.0707	0.8854	106
Panel B: Rolling mode					
EW_ROLL	0.7772	0.9029	0.1964	4.9776	36
AR1_ROLL	0.5301	0.6435	0.1126	3.6833	36
DMA_ROLL	0.2739	0.2244	0.0558	1.0527	36
DMS_ROLL	0.3448	0.2583	0.0928	1.2793	36
BMA_ROLL	0.5867	0.2539	0.1170	1.1630	36
BMS_ROLL	0.5817	0.2467	0.1296	1.2134	36
ITMA_ROLL	0.3367	0.2097	0.0710	0.8849	36

Table 4
Results of the MCS test for different forecasting models (Beijing).

Model	MSFE						MAFE					
	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D
Panel A: Recursive mode												
EW	0.0057	0.0001	0.0065	0.0055	0.0003	0.0113	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR1	0.0188	0.0078	0.0771	0.0135	0.0112	0.0247	0.0003	0.0004	0.0005	0.0000	0.0001	0.0003
DMA	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
DMS	0.2558*	0.3160*	0.2603*	0.2558*	0.3183*	0.2603*	0.0451	0.0451	0.0392	0.0435	0.0461	0.0392
BMA	0.0057	0.0000	0.0007	0.0020	0.0000	0.0055	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BMS	0.0188	0.0086	0.0587	0.0099	0.0112	0.0113	0.0003	0.0001	0.0020	0.0001	0.0001	0.0005
ITMA	0.0188	0.0000	0.0020	0.0099	0.0000	0.0113	0.0020	0.0016	0.0000	0.0000	0.0000	0.0003
Panel B: Rolling mode												
EW_ROLL	0.0820	0.1275*	0.3471*	0.2149*	0.3085*	0.1655*	0.0343	0.0650	0.1267*	0.0997	0.1794*	0.0695
AR1_ROLL	0.4870*	0.5455*	0.5352*	0.3769*	0.4171*	0.5675*	0.2104*	0.3288*	0.2159*	0.3841*	0.4239*	0.2393*
DMA_ROLL	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
DMS_ROLL	0.4870*	0.5455*	0.5352*	0.3769*	0.4171*	0.5675*	0.2677*	0.3288*	0.2159*	0.3841*	0.4239*	0.2728*
BMA_ROLL	0.4870*	0.5455*	0.5352*	0.3769*	0.4171*	0.5675*	0.3507*	0.3507*	0.3579*	0.3841*	0.4239*	0.3579*
BMS_ROLL	0.0820	0.1096*	0.4858*	0.2149*	0.3085*	0.3303*	0.0343	0.0546	0.1772*	0.0997	0.1794*	0.1112*
ITMA_ROLL	0.0820	0.0438	0.0419	0.0329	0.0854	0.0230	0.0343	0.0120	0.0042	0.0014	0.0178	0.0068

Note: * indicates a p-value larger than 0.1, which means that the corresponding model survives the MCS test under a specific loss function and test statistic. The bold p-value 1.000 indicates a model that performs better than the other models.

(MSFE) as defined in Eq. (14) for all of the forecasting models. This table shows the summary of the forecasting errors for all 30 selected major cities in China.

Table 3 shows that with regard to recursive forecasting, first, the average MSFE for DMA is 0.2759, which is the smallest result among the six forecasting models. In addition, the forecasting deviation of DMA is 0.2258, which is also quite tiny compared to that of the other six models. This result indicates the forecasting robustness of DMA for the different city samples. Third, the second-best-performing model is BMS, with an average MSFE of 0.2833 and the second smallest forecasting deviation of 0.2200. This result implies that in some cases, the traditional Bayesian forecasting model without the “forgetting effect” (like the DMA model) can also be very useful. Fourth, although ITMA does not get the best accurate forecasts, it obtains the smallest forecasting deviation of 0.2099, implying that this method may get robust forecasts for different city samples. Last, the EW and AR1 models are not able to give satisfactory forecasting results.

With regard to rolling-window forecasting, firstly the DMA produces not only the smallest average MSFE of 0.2739, but also the second minimum forecasting deviation of 0.2244. This result indicates that the DMA method not only improves forecasting accuracy significantly, but it also achieves stable forecasting precision among different samples. Secondly, however, the BMS model does not behave as well as it does in the recursive mode, which implies this model's instability in forecasting accuracy with different data samples. Thirdly, ITMA again achieve the smallest forecasting deviation of 0.2097, indicating its advantage of forecasting robustness across different city samples. Finally, the traditional EW and AR1 models also perform poorly with the rolling-window mode.

It is worth noting that rigorous statistical analysis is necessary to get robust evaluation results for different forecasting models. In the following procedures, we use an advanced statistical test, i.e., the MCS test as proposed by Hansen et al. (2011), to choose the superior models. As discussed above, the models with p-values larger than 0.1 can be treated as “survival models,” as larger p-values indicate better forecasting accuracy. Tables 4–6 report the p-values for MCS tests under different loss functions and test statistics for the cities of Beijing, Haikou and Hohhot. The latter two cities have the highest and the

Table 5
Results of the MCS test for different forecasting models (Haikou).

Model	MSFE						MAFE					
	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D
Panel A: Recursive mode												
EW	0.0806	0.1075*	0.1171*	0.0049	0.0107	0.1014*	0.0016	0.0024	0.0072	0.0000	0.0000	0.0031
AR1	0.0986*	0.1202*	0.1171*	0.1363*	0.1633*	0.1114*	0.0265	0.0185	0.0175	0.0207	0.0319	0.0273
DMA	0.2840*	0.2840*	0.2839*	0.2922*	0.2946*	0.2839*	0.4289*	0.4289*	0.4372*	0.4302*	0.4320*	0.4372*
DMS	0.1454*	0.1464*	0.1171*	0.1604*	0.1683*	0.1208*	0.0814*	0.0936	0.1035*	0.1237*	0.1314*	0.0938*
BMA	0.1454*	0.1202*	0.1171*	0.1363*	0.1633*	0.1114*	0.0322	0.0185	0.0172	0.0207	0.0319	0.0229
BMS	0.1133*	0.1202*	0.1171*	0.1363*	0.1633*	0.1114*	0.0244	0.0153	0.0172	0.0207	0.0319	0.0202
ITMA	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
Panel B: Rolling mode												
EW_ROLL	0.0004	0.0011	0.0001	0.0000	0.0005	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR1_ROLL	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	0.7474*	0.6300*	0.6338*	0.6165*	0.6196*	0.6338*
DMA_ROLL	0.4983*	0.3772*	0.3871*	0.3649*	0.3711*	0.3932*	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
DMS_ROLL	0.4983*	0.3202*	0.3711*	0.1688*	0.1928*	0.3932*	0.6814*	0.2730*	0.2540*	0.0620	0.1110*	0.3068*
BMA_ROLL	0.4983*	0.1565*	0.3711*	0.0234	0.0453	0.3932*	0.7474*	0.3893*	0.3921*	0.1125*	0.1353*	0.5814*
BMS_ROLL	0.4983*	0.1409*	0.2104*	0.0033	0.0207	0.3503*	0.2612*	0.1535*	0.1297*	0.0620	0.1110*	0.2591*
ITMA_ROLL	0.4983*	0.3202*	0.2104*	0.0036	0.0207	0.3932*	0.2612*	0.2209*	0.0677	0.0015	0.0134*	0.2591*

Note: * indicates a p-value larger than 0.1, which means that the corresponding model survives the MCS test under a specific loss function and test statistic. The bold p-value 1.000 indicates that the corresponding model performs better than the other models.

Table 6
Results of the MCS test for different forecasting models (Hohhot).

Model	MSFE						MAFE					
	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D
Panel A: Recursive mode												
EW	0.0029	0.0008	0.0008	0.0003	0.0010	0.0019	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR1	0.0304	0.0163	0.0017	0.0023	0.0044	0.0096	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
DMA	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
DMS	0.0342	0.0294	0.0067	0.0051	0.0068	0.0199	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
BMA	0.0029	0.0002	0.0007	0.0001	0.0004	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BMS	0.0029	0.0055	0.0017	0.0023	0.0044	0.0054	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ITMA	0.1678*	0.1678*	0.1708*	0.0905	0.0934	0.1708*	0.0662	0.0662	0.0745	0.0298	0.0320	0.0745
Panel B: Rolling mode												
EW_ROLL	0.0165	0.0290	0.0050	0.0011	0.0153	0.0286	0.0028	0.0036	0.0012	0.0000	0.0014	0.0053
AR1_ROLL	0.1022*	0.1462*	0.0051	0.0144	0.0369	0.0382	0.1118*	0.1460	0.0030	0.0252	0.0500	0.0269
DMA_ROLL	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
DMS_ROLL	0.1022*	0.1462*	0.0420	0.0127	0.0369	0.0775	0.1118*	0.0909	0.0030	0.0200	0.0500	0.0269
BMA_ROLL	0.1413*	0.1462*	0.0553	0.0341	0.0412	0.0775	0.2976*	0.2405*	0.2023*	0.1763*	0.2004*	0.3248*
BMS_ROLL	0.1022*	0.0534	0.0051	0.0005	0.0412	0.0382	0.1118*	0.0254	0.0026	0.0001	0.0024	0.0161
ITMA_ROLL	0.1413*	0.1462*	0.0143	0.0144	0.0369	0.0382	0.2976*	0.2405*	0.2023*	0.1763*	0.2004*	0.3248*

Note: * indicates a p-value larger than 0.1, which means that the corresponding model survives the MCS test under a specific loss function and test statistic. The bold p-value 1.000 indicates that the corresponding model performs better than the other models.

lowest growth rates in house prices (respectively) during the data sample period. Table 7 presents the results of the MCS test for the selected 30 major cities.

Tables 4 to 6 indicate that for house price forecasting in any individual city, several models can survive the MCS test. Taking Beijing as a first example, Table 4 shows that with the recursive mode, only DMA model can survive under both the MSFE and MAFE criteria. However, for the rolling mode, except ITMA, almost all of the other six models can pass the MCS test and obtain acceptable forecasting accuracy. The overall results demonstrate that DMA can achieve the best forecasting performance.

When forecasting the house prices in Haikou (which has the highest average rate of growth in house price), Panel A of Table 5 shows that almost all of the seven models can survive under the MSFE criterion, but only DMA, DMS and ITMA can pass the MCS test under the MAFE criterion. In general, DMA, DMS and ITMA are survival models under all criteria and test statistics, in which ITMA seems to perform better than the others. Unlike in the situation of Beijing, Panel B of Table 5 indicates that when the rolling forecasting results are tested, it is the AR1 model (and not DMA or ITMA) that performs best under MSFE. When MAFE is considered, DMA again achieves the best forecasting accuracy.

When we consider the city of Hohhot (which has the lowest house price growth rate in China), Panel A of Table 6 shows that only DMA can survive the MCS test under all criteria and test statistics in recursive forecasting mode, and in some cases ITMA is acceptable under the MSFE criterion. Panel B indicates that, except EW model, almost all of the other six models can partly produce acceptable accuracy in rolling forecasting.

In general, the three individual cities discussed here present no consistent proof concerning which model is superior in forecasting house prices. Thus, to get an overview of the performance of different models, we pool the forecasting results for all of the selected 30 cities, and again use MCS to test the overall forecasting accuracy of the models. The empirical results shown in Table 7 are quite simple. With regard to recursive forecasting, Panel A of Table 7 shows that only DMA can survive the MCS test under all of the loss functions and test statistics. All the MCS p-values of DMA equaling to one indicates that it performs very well in forecasting house price in major Chinese cities. While in the situation of rolling forecasting, the empirical results are a bit more complicated. Among all the seven forecasting models, not only DMA, but BMA and ITMA can survive the MCS tests. Under the criterion of MSFE, DMA again obtain the best forecasting accuracy. While ITMA is the superior model to the others with respect to MAFE criterion.

However, it is worth noting that there are only 36 monthly rolling forecasting observations in our empirical test, and there are 106 monthly recursive forecasting observations. Thus, with respect to long-run performance, the testing results for recursive forecasting in this empirical study are more reliable. In short, no matter what forecasting modes or forecasting horizons are used, DMA performs the best among all of the models considered here. In some cases, ITMA proposed by Kapetanios et al. (2008) can also achieves quite satisfactory forecasting accuracy.

4.2. Average number of predictors

Compared to traditional prediction models, an important advantage of DMA is that it allows explanatory variables (predictors) to change over time. This means that the numbers and types of explanatory variables at different time points may change according to their past forecasting performance. Koop and Korobilis (2012) point out that the average (or expected) number of explanatory variables (predictors) used by DMA at time t is

$$E(size_t) = \sum_{k=1}^{K} \pi_{(t|t-1,k)} size_{(k)}, \tag{20}$$

where $size_{(k)}$ is the number of explanatory variables in the kth model. Fig. 2 presents the median of the $E(size_t)$ for the 30 major cities in China. The 1/4 and 3/4 quantiles of $E(size_t)$ are also indicated in this figure. Fig. 2 shows that as time passes, the average number of explanatory variables decreases. At the beginning of 2007, about 8 predictors are used, but this number falls to about 5.5 by the end of 2015. Considering that a total of 12 alternative predictors are tested in this paper, these results imply that DMA can effectively select good predictors and eliminate bad predictors from the optional variables. As time goes on, more forecasting performance information is collected by DMA. Thus, the redundant predictors with poor forecasting accuracy are eliminated, and good predictors are retained in the DMA process. As fewer predictors are adopted over time, the calculation burden of DMA is diminished, and the forecasting effectiveness is improved.

Table 7
Results of the MCS test for different forecasting models (summary of 30 major cities in China).

Model	MSFE						MAFE					
	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D	T_R	T_{SQ}	T_{Max}	T_Q	T_F	T_D
Panel A: Recursive mode												
EW	0.0010	0.0025	0.0024	0.0000	0.0000	0.0031	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR1	0.0100	0.0063	0.0065	0.0001	0.0001	0.0159	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
DMA	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***
DMS	0.0151	0.0083	0.0065	0.0003	0.0003	0.0212	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BMA	0.0100	0.0063	0.0065	0.0001	0.0000	0.0120	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BMS	0.0100	0.0063	0.0065	0.0001	0.0000	0.0146	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ITMA	0.0768	0.0768	0.0737	0.0748	0.0749	0.0737	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Panel B: Rolling mode												
EW_ROLL	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR1_ROLL	0.0004	0.0026	0.0012	0.0003	0.0003	0.0016	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
DMA_ROLL	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	0.4395*	0.5140*	0.4201*	0.5247*	0.5252*	0.3765*
DMS_ROLL	0.0109	0.0510	0.0388	0.0244	0.0250	0.0608	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BMA_ROLL	0.4928*	0.4368*	0.3472*	0.4437*	0.3523*	0.4178*	0.4395*	0.5140*	0.4183*	0.5247*	0.5252*	0.3724*
BMS_ROLL	0.0000	0.0000	0.0002	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ITMA_ROLL	0.4928*	0.4368*	0.3326*	0.4437*	0.4443*	0.4178*	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***	**1.0000***

Note: * indicates a p-value larger than 0.1, which means that the corresponding model survives the MCS test under a specific loss function and test statistic. The bold p-value 1.000 indicates that the corresponding model performs better than other models.

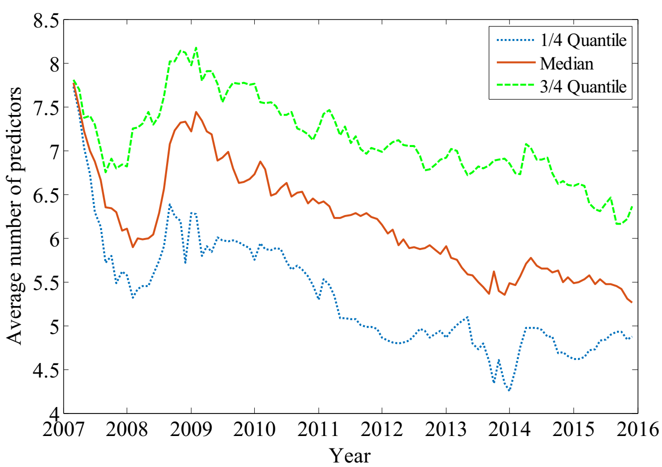

Fig. 2. Average number of predictors used in DMA recursive forecasting for 30 major cities in China.

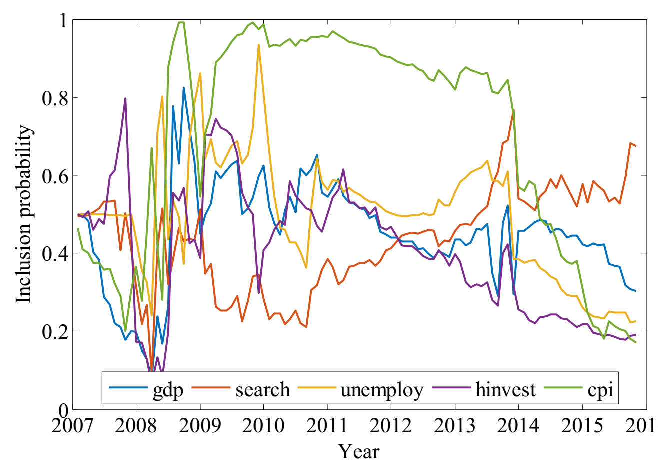

Fig. 3. Inclusion probabilities for predictors in recursive DMA forecasting for 30 major cities in China.

4.3. Inclusion probability of predictors in DMA forecasting

The probability of inclusion for a predictor (e.g., the growth rate of CPI, GDP or Google search index) refers to the sum of probabilities ($\pi_{(t|t-1,k)}$) that a given predictor will be included among the set of viable predictors in the forecasting model k (k=1, 2, ..., K) of DMA at time t. In other words, the higher the probability of inclusion for a predictor, the more forecasting weight is assigned to that variable, and the higher forecasting power that predictor has. To illustrate this approach clearly, Fig. 3 shows the probabilities for five typical predictors in DMA forecasting.

As Fig. 3 indicates, all of the five predictors considered here present volatile forecasting probabilities over the whole data sample period. At the beginning of the forecasting sample (in 2007), investment in the real estate market (hinvest) shows higher forecasting power than the other predictors. The forecasting power of hinvest also increases over time, as other predictors show decreasing probabilities. During the middle part of the sample period (from 2008 to 2014), the CPI is obviously superior to other predictors, and unemployment performs also quite well most of the time. However, there is clear evidence that from 2009 to 2014, the forecasting power of the Google search index (search) increases significantly. In stark contrast, GDP, unemployment, hinvest and CPI gradually lose their forecasting power. At the end of sample period (from year 2014 to 2015), the Google search index overtakes the others to become the best predictor. In addition, the forecasting probability of the Google search index continually rises, and the probabilities of other predictors fall remarkably.

According to economic theory, economic variables such as GDP, CPI or investment should have more ability to explain changes in house prices. However, we find that the explanatory ability of these traditional macroeconomic predictors has decreased gradually in recent years, and the Google search index is gaining ever more forecasting power. There are two possible reasons for this trend. First, in 2008 the U.S. subprime mortgage crisis swept through the global economy, and the effects of this event are still being felt. Also since 2008, the Chinese government and central bank have implemented many credit policies to stimulate or reduce supply and demand in the real estate market. Thus, house prices have fluctuated more due to uncertainty over the government's policies than in response to macroeconomic indicators. Second, according to a report by the China Internet Network Information Center (CNNIC), the number of Chinese Internet users increased from 111 million in 2005 to 688 million in 2015, and 71.5% of these people are accessing the Internet with mobile devices. Due to uncertainty over the trends in China's economic development and real estate market policy, more and more house seekers are using the

Internet to find information for their purchase decisions. Therefore, changes in the search indexes on house prices contain more information about the real demand from house purchasers, and thus a search index may have more ability to explain house price changes in China.

5. Conclusions

This paper adopts the dynamic model averaging (DMA) method to predict the growth rate in house prices for 30 major cities in China. Unlike previous empirical studies, our paper uses both recursive and rolling forecasting modes in DMA. We also apply a rigorous statistical method, namely the model confidence set (MCS) test, to compare the forecasting performance of DMA with that of other models. Last, we use the Google search index as an additional predictor beyond the traditional economic variables to forecast changes in Chinese house prices.

The main empirical results show that compared with the traditional time series models and other model averaging approaches, DMA is more flexible and effective, as it allows both the models and the coefficients to change over time. Under the criteria of MSFE, MAFE and many different test statistics, the MCS test indicates that DMA achieves significantly higher forecasting accuracy than other models in both the recursive and rolling forecasting modes. However, no single predictor is found to have absolute superiority over others. The best predictors for Chinese house prices vary greatly over time. Also, the Google search index for house prices has surpassed the forecasting ability of traditional macroeconomic variables in recent years.

Acknowledgements

We would like to show our sincere gratitude to an anonymous reviewer, as well as Professor Sushanta Mallick, Co-Editor-in-Chief of Economic Modelling, whose comments and suggestions greatly improved the quality of this paper. The authors are also grateful for the financial support from the National Natural Science Foundation of China (71371157 and 71671145), the Humanities and Social Science Fund of Ministry of Education (14YJC790073, 15YJA790031 and 16YJA790062), and the Young Scholar Fund of Science & Technology Department of Sichuan Province (2015JQO010), fund of high-level social science research team of Sichuan Province, Fundamental Research Funds for the Central Universities (26816WCX02).

References

Akinsomi, O., Aye, G.C., Babalos, V., Economou, F., Gupta, R., 2016. Real estate returns predictability revisited: novel evidence from the US REITs market. Empir. Econ. 51, 1165–1190.

Bork, L., Møller, S.V., 2015. Forecasting house prices in the 50 states using dynamic model averaging and dynamic model selection. Int. J. Forecast. 31, 63–78.

Brown, J.P., Song, H., McGillivray, A., 1997. Forecasting UK house prices: a time varying coefficient approach. Econ. Model. 14, 529–548.

Case, K.E., Shiller, R.J., 1990. Forecasting prices and excess returns in the housing market. AREUA J. 18, 253–273.

Crawford, G.W., Fratantoni, M.C., 2003. Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. Real Estate Econ. 31, 223–243.

Das, P., Ziobrowski, A., Coulson, N.E., 2015. Online information search, market fundamentals and apartment real estate. J. Real Estate Financ. Econ. 51, 480–502.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13, 253–263.

DiPasquale, D., Wheaton, W.C., 1994. Housing market dynamics and the future of housing prices. J. Urban Econ. 35, 1–27.

Fortescue, T.R., Kershenbaum, L.S., Ydstie, B.E., 1981. Implementation of self-tuning regulators with variable forgetting factors. Automatica 17, 831–835.

Ghysels, E., Plazzi, A., Valkanov, R., Torous, W., 2013. Forecasting real estate prices. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting 2. Elsevier, Amsterdam, 509–580.

Hadavandi, E., Ghanbari, A., Mohsen Mirjani, S., Abbasian, S., 2011. An econometric panel data-based approach for housing price forecasting in Iran. Int. J. Hous. Mark. Anal. 4, 70–83.

Hamilton, B.W., Schwab, R.M., 1985. Expected appreciation in urban housing markets. J. Urban Econ. 18, 103–118.

Hansen, P.R., 2005. A test for superior predictive ability. J. Bus. Econ. Stat. 23, 365–380.

Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79, 453–497.

Kapetanios, G., Labhard, V., Price, S., 2008. Forecasting using Bayesian and information-theoretic model averaging: an application to UK inflation. J. Bus. Econ. Stat. 26, 33–41.

Koop, G., Korobilis, D., 2011. UK macroeconomic forecasting with many predictors: which models forecast best and when do they do so? Econ. Model. 28, 2307–2318.

Koop, G., Korobilis, D., 2012. Forecasting inflation using dynamic model averaging. Int. Econ. Rev. 53, 867–886.

Koop, G., Leon-Gonzalez, R., Strachan, R., 2009. On the evolution of the monetary policy transmission mechanism. J. Econ. Dyn. Control 33, 997–1017.

Malpezzi, S., 1999. A simple error correction model of house prices. J. Hous. Econ. 8, 27–62.

Man, G., 2015. Competition and the growth of nations: international evidence from Bayesian model averaging. Econ. Model. 51, 491–501.

Nneji, O., Brooks, C., Ward, C.W.R., 2013. House price dynamics and their reaction to macroeconomic changes. Econ. Model. 32, 172–178.

Piegorsch, W.W., An, L., Wickens, A.A., Webster West, R., Peña, E.A., Wu, W., 2013. Information-theoretic model-averaged benchmark dose analysis in environmental risk assessment. Environmetrics 24, 143–157.

Plakandaras, V., Gupta, R., Gogas, P., Papadimitriou, T., 2015. Forecasting the U.S. real house price index. Econ. Model. 45, 259–267.

Primiceri, G., 2005. Time varying structural vector autoregressions and monetary policy. Rev. Econ. Stud. 72, 821–852.

Próchniak, M., Witkowski, B., 2013. Time stability of the beta convergence among EU countries: bayesian model averaging perspective. Econ. Model. 30, 322–333.

Raftery, A.E., Karny, M., Ettler, P., 2010. Online prediction under model uncertainty via dynamic model averaging: application to cold rolling mill. Technometrics 52, 52–56.

Rapach, D.E., Strauss, J.K., 2009. Differences in housing price forecast ability across U.S. states. Int. J. Forecast. 25, 351–372.

West, K.D., 1996. Asymptotic inference about predictive ability. Econometrica 64, 1067–1084.

White, H., 2000. A reality check for data snooping. Econometrica 68, 1097–1126.