University of Pretoria
Department of Economics Working Paper Series

The Role of Investor Sentiment in Forecasting Housing Returns in China: A Machine Learning Approach

Oguzhan Cepni
Central Bank of the Republic of Turkey
Rangan Gupta
University of Pretoria
Yigit Onay
Central Bank of the Republic of Turkey
Working Paper: 2020-55
June 2020

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

The Role of Investor Sentiment in Forecasting Housing Returns in China: A Machine Learning Approach

Oguzhan Cepni [a], Rangan Gupta[b], Yigit Onay[a]

Abstract

This paper analyzes the predictive ability of aggregate and dis-aggregate proxies of investor sentiment, over and above standard macroeconomic predictors, in forecasting housing returns in China, using an array of machine learning models. Using a monthly out-of-sample period of 2011:01 to 2018:12, given an in-sample of 2006:01-2010:12, we find that indeed the new aligned investor sentiment index proposed in this paper has greater predictive power for housing returns than the a principal component analysis (PCA)-based sentiment index, used earlier in the literature. Moreover, shrinkage models utilizing the dis-aggregate sentiment proxies do not result in forecast improvement indicating that aligned sentiment index optimally exploits information in the dis-aggregate proxies of investor sentiment. Furthermore, when we let the machine learning models to choose from all key control variables and the aligned sentiment index, the forecasting accuracy is improved at all forecasting horizons, rather than just the short-run as witnessed under standard predictive regressions. This result suggests that machine learning methods are flexible enough to capture both structural change and time-varying information in a set of predictors simultaneously to forecast housing returns of China in a precise manner. Given the role of the real estate market in China's economic growth, our result of accurate forecasting of housing returns, based on investor sentiment and macroeconomic variables using state-of-the-art machine learning methods, has important implications for both investors and policymakers.

Keywords: Housing prices; Investor sentiment; Bayesian shrinkage; Time-varying parameter model
JEL classification: C22; C32; C52; G12; R31

[a] Central Bank of the Republic of Turkey, Haci Bayram Mah. Istiklal Cad. No:10 06050, Ankara, Turkey; Email addresses: Oguzhan.Cepni@tcmb.gov.tr; Yigit.Onay@tcmb.gov.tr

[b] Department of Economics, University of Pretoria, Pretoria, 0002, South Africa; Email address: rangan.gupta@up.ac.za

1. Introduction

Financialization of the housing market, i.e., its treatment as a commodity, though a contentious issue due to housing's primary role to serve as a social good, is now a well-established global fact (Aalbers, 2016). And in this regard, China, a major player in the world economic system, is not far behind, though the process started more recently compared to the western world (Wu et al., 2020). The commodification of the housing market is considered to be one of the major drivers of China's economic development (Hsing, 2010; Tao et al., 2010; Lin, 2014; Wu, 2015), and is believed to have played an important role in crisis management. For instance, the suspension of welfare housing provision in 1998 as a response to the 1997 Asian financial crisis (Logan et al., 2010), and more recently, the housing boom triggered by the injection of 4 trillion yuan into infrastructure and urban development after the 2008 global financial crisis (Deng and Chen, 2019). Put alternatively, financialization of the housing market in China (and around the world) basically implies that price co-movement between real and financial assets, i.e., the equity market (as well as between different assets), is likely to increase (Hong and Li, 2019a), with the underlying driver being investor sentiment (Tang and Xiong, 2012). Moreover, while documenting the relatively large amount of existing evidence of investor sentiment in affecting Chinese stock returns, Su et al., (2020) point out that, since individual investors account for a large proportion of Chinese stock investors, they not only tend to make irrational trades in the stock market, but are also more likely to carry their emotions to other markets.

Now, the importance of the real estate markets for China's past and continued economic growth, and the key overall role of house prices as a leading indicator of the macro-economy is well-recognized (Chow et al., 2018). This is especially due to the introduction of neo-liberal reforms in the 1990s, and particularly since 1998, when public sector housing allocation was replaced by market allocation and quasi-privatization of property (Theurillat et al., 2016). Hence, accurate forecasting of housing returns based on the information content of investor sentiment in the wake of financialization is an important question for policy authorities to gauge the future path of the overall domestic economy. Given China's position in the global economy as the second-largest economy (after the United States (US)), with its share of global gross domestic product (GDP) adjusted for purchasing-power-parity (PPP) being 19.72% (Global Competitiveness Report, 2019), performance of the Chinese economy is also a pertinent issue for policymakers around the world.

Against this backdrop, the objective of our paper is to forecast composite housing returns for 70 large and medium-sized cities in China over the monthly period of 2011:01 to 2018:12, given an in-sample period of 2006:01 to 2010:12, based on investor sentiment (controlling for other predictors) using a variety of machine learning methods (such as generalized approximate message passing (GAMP), Bayesian model averaging, Ridge regression, least absolute shrinkage operator (LASSO), Elastic Net). These shrinkage-based approaches allow us to efficiently conduct the forecasting experiment, without suffering from the "curse of dimensionality", especially in the context of a time-varying framework with multiple predictors and relative short-span (13 years) of data (in our case 156 monthly observations). Our paper can be considered to be an extension of Hong and Li (2019b), whereby they use wavelet analysis to provide in-sample evidence of the predictability of housing returns in China due to investor sentiment. However, since in-sample predictability does not guarantee forecasting gains, and as pointed out by Campbell (2008), and Bork and Møller (2015) specifically for housing returns, that the ultimate test of any predictive model (in terms of the econometric methodologies and the predictors used) is in its out-of-sample performance, evidence of forecastability, if it exists, would provide more robust evidence (relative to an in-sample analysis) of the role of investor sentiment for future housing returns.

One must recall that investor sentiment is a latent variable, and needs to be derived from appropriate proxies. Given this Hong and Li (2019b), followed Baker and Wurgler (2006) to form the investor sentiment index using the principal component analysis (PCA) to aggregate the information from six individual proxies (the closed-end fund discount; average first-day returns on initial public offerings (IPOs); the ratio of the number of advancing stocks to the number of declining stocks; new A-share market accounts; market turnover rate; and consumer confidence index), which we use as well, both as an index and individually. But as an alternative to PCA, we also use partial least squares (PLS) to construct the sentiment index. Econometrically speaking, the first principal component, is indeed the best combination of the six proxies that represents the highest percentage of the total variations of the proxies. Since all the proxies may have approximation errors to the true but unobservable investor sentiment, and these errors are parts of their variations, the first principal component can potentially contain a substantial amount of common approximation errors that are possibly not relevant for forecasting asset returns (Boivin and Ng, 2006; Bai and Ng, 2008). Given this, we align the investor sentiment measure with the purpose of explaining the housing returns by extracting the most relevant common component from the proxies. In other words, we separate out information in the proxies that is relevant to the expected housing returns from the error or noise, by using the partial least squares (PLS) method originally developed by Wold (1966, 1975), and applied to financial and housing returns data more recently by Kelly and Pruitt (2013, 2015), Bork and Møller (2018), Bork et al., (2020), and Cepni et al., (2020a). Hence, the usage of the PLS to obtain an aligned investor sentiment index, can also be considered as a contribution of our study, besides the full-fledged out-of-sample forecasting exercise using machine learning methods.

To the best of our knowledge, this is the first attempt to forecast housing returns in China using aggregated and dis-aggregated proxies of investor sentiment and other macroeconomic and financial variables, based on a wide array of machine learning methods. The two papers that we could find which have produced out-of-sample forecasting of housing returns for China is that of Wei and Cao (2017), and Salisu and Gupta (forthcoming). While the latter paper shows that monthly dis-aggregated oil shocks, i.e., supply, global economic activity, oil-specific demand, and oil inventory demand, can be used to forecast quarterly housing returns of China based on a mixed-frequency model, the former paper highlights the role of a Google search index (associated with city name plus house price), instead of fundamental macroeconomic or monetary indicators, based on a dynamic model averaging (DMA) framework.

The remainder of the paper is organized as follows: Section 2 presents the data; Section 3 outlines the methodologies used, with Section 4 presenting the main econometric results along with robustness analysis, and Section 5 concludes the paper.

2. Data

Housing price is downloaded for 70 large and medium-sized cities in China over the monthly period of 2006:01 to 2018:12.[1] Then, the composite housing price index is computed as the average of these indices. The monthly housing price return calculated as using the formula: $Houret_t = ln(HPI_t) - ln(HPI_{t-1}) \times 100$ where HPI_t is the monthly house price index at time t. We construct our sentiment index using the six individual sentiment proxies based on the work of Hong and Li (2019b)[2]. Following Baker and Wurgler (2006), they form the investor sentiment index using PCA (Investor.Sent.PCA) to aggregate the information from six individual proxies: the closed-end fund discount (Dcef), average first-day returns on IPOs (RIpo), ratio of the number of advancing stocks to the number of declining stocks (Adrt), new A-share market accounts (NA), market turnover rate (Turn), and consumer confidence index (CCI)[3]. As an alternative to the PCA, we use partial least squares to construct the sentiment index. Finally, we also collect a set of key economic variables which are growth of industrial production (IP), consumer price index inflation (CPI), the People's Bank of China's policy rate (IR), and returns of the Shanghai composite stock market index (SMR). Raw values of all control variables are downloaded from the Bloomberg terminal.

3. Methodologies

3.1. The construction of a new sentiment index

To construct our investor sentiment index (Investor.Sent.PLS) using the information contained in each of six individual sentiment proxies, we employ the PLS method to the same six proxies. In particular, we utilize the PLS method using Friedman et al.,'s, (2001) two-step approach. The algorithm starts by standardizing each proxies x_j ($j = 1, \dots, p$) to have a zero mean and unit variance. Then, uni-variate regression coefficients $\widehat{\gamma_{1j}} = \langle x_j, y \rangle$ are computed for each j. From this, the first PLS direction $z_1 = \sum_j \widehat{\gamma_{1j}} x_j$ is obtained as the weighted sum of the vector of uni-variate regression coefficients and the original set of sentiment proxies. Hence, the construction of the PLS direction takes into account the degree of association between housing returns and common factors. In the following step, the "target" variable y is regressed on z_1, resulting in a coefficient θ_1, and then all inputs are orthogonalized with respect to z_1. This process is iterated until PLS produces a sequence of $l < p$ orthogonal directions.

Since PLS utilizes the housing returns to construct the directions, its solution path is a non-linear function of housing returns. While PCA finds directions that maximize only the variance, PLS aims for the

[1]Data can be downloaded from the official website of China's National Bureau of Statistics: http://data.stats.gov.cn/.

[2]The data of sentiment proxies ends on December, 2018. As a result, the sample period used in our analysis ends on that date.

[3]We thank Dr. Yun Hong and Dr. Yi Li for sharing their data in this regard.

directions that have high variance and high correlation with the target variable which intuitively could increase the forecasting power of a PLS-based index compared to a PCA-based index.

More specifically, the m^{th} PLS direction $\widehat{\gamma_m}$ solves the following optimization problem:

$$\begin{aligned} &\max_{\alpha} && Corr^2(y, X_\alpha)Var(X_\alpha), \\ &\text{subject to} && \|\alpha\| = 1, \quad \alpha' S \widehat{\gamma_l} = 0, \quad l = 1, ..., m-1 \end{aligned} \tag{1}$$

where S represents the sample covariance matrix of the x_j. We choose the first common component as a new investor sentiment index, which efficiently incorporates all the relevant information from the each of the six sentiment proxies for housing returns.

3.2. Time varying parameter regressions : machine learning approaches

After the construction of new sentiment index, this section introduces a comprehensive list of competing specifications and estimation algorithms which are presented in the following subsections. All models are estimated on an expanding window using only information available at the time of forecast.

3.2.1. Generalized approximate message passing (GAMP) algorithm

Although the Bayesian approach using Markov Chain Monte Carlo Methods (MCMC) is a powerful tool to take into account the changing nature of the relationship between variables, computational concerns with these methods as well as large errors related to repeated sampling through Monte Carlo, make it harder to rely on them in case the dimension of the econometric model is high (Angelina et al., 2016). As an alternative to computational limitations, message passing algorithms come to the forefront representing a highly efficient and easy-to-implement Bayesian estimation algorithm which makes it possible to take stochastic volatility and parameter instability into account with a large set of predictors. Moreover, unlike ”well-established” MCMC algorithms, GAMP-based algorithms require minimal or no tuning.

TVP-GAMP offers an efficient way of estimation by allowing time-varying parameter and variable selection together with stochastic volatility simultaneously with no restrictions on the number of predictors. TVP-GAMP procedure relies on the visualization of the relation between random variables within the framework of factor graphs.[4] By factorizing joint posterior distribution functions of random variables into smaller parts, the procedure offers a localised way to iteratively approximate complex set of conditional marginal distributions with an approach called the sum-product algorithm, which is also known as “Belief Propagation”.[5]

To illustrate the estimation process through GAMP procedure, a time-varying parameter specification with stochastic volatility is given as follows:

$$y_t = x_t\beta_t + \varepsilon_t \tag{2}$$

[4]Factor graph approach decomposes random variables into quantities of lower dimensions. See Korobilis (2019) for simplified illustration of the factor graph representation.

[5]See for example, Pearl (1982) for details.

subject to an initial condition for β_t at $t = 0$, where y_t is the t^{th} observation on the dependent variable, $t = 1, \ldots, T$; x_t is a $1 \times q$ vector of predictors, β_t is a $q \times 1$ vector of coefficients, and $\varepsilon_t \sim N\left(0, \sigma_t^2\right)$ with σ_t^2 the time-varying variance parameter.

Equivalently, rewriting the regression in the static form, we have the matrix representation as follows:

$$y = \mathcal{X}\beta + \varepsilon \tag{3}$$

where $y = [y_1, \ldots, y_T]'$ and $\varepsilon = [\varepsilon_1, \ldots, \varepsilon_T]'$ are column vectors representing the observations y_t and ε_t respectively, $\beta = \left[\beta', \beta_1', \ldots, \beta_T'\right]'$ is a $(T+1)q \times 1$ vector containing the coefficients. The number of parameters to be estimated for the coefficients is equal to $q = (T+1)p$ which is not possible to estimate with the classical OLS procedure. State space models have been offered in the literature to overcome the problem of identification such that a stochastic process, most typically random walk, is assumed for the coefficients and the estimation is accomplished through Markov chain Monte Carlo methods for these models.

As an alternative, TVP-GAMP attempts to identify the whole set of coefficients in the TVP regression with data driven hierarchical shrinkage priors. Consider i.i.d prior $p(\beta) = \prod_{i=1}^{q} p(\beta_i)$, [6] then the marginal posterior for $\beta_i, i = 1, \ldots, q$ obtained through Bayes Theorem, requires evaluation of a $(q-1)$ -dimensional integral of the form

$$\begin{aligned} p(\beta_i|y) &= \int p(\beta|y) d\beta_{j\neq i} \\ &= \int p(y|\beta) p(\beta) d\beta_{j\neq i} \\ &= p(\beta_i) \int p(y|\beta) \prod_{j=1, j\neq i}^{q} p\left(\beta_j\right) d\beta_{j\neq i} \end{aligned} \tag{4}$$

Computation of the above summation can be quite cumbersome especially in case of high dimensionality problem in parameters. Factorising the above-mentioned posterior distribution through factor graphs, we define $\mu_{p(\bullet)\to a}$ the message passed from probability function $p(\bullet)$ to random variable a, then:

$$p(\beta_i|y) = \mu_{p(\beta_i)\to\beta_i} \prod_{t=1}^{T} \mu_{p(y_t|\beta)\to\beta_i} \tag{5}$$

where $\mu_{p(\beta_i)\to\beta_i} = p(\beta_i)$. According to sum-product rule we further have

$$\mu_{p(y_t|\beta)\to\beta_i} = \int p(y_t|\beta) \prod_{j=1, j\neq i}^{p} \mu_{\beta_j \to p(y_t|\beta)} d\beta_{j\neq i} \tag{6}$$

[6] Sparse Bayesian learning (SBL) prior as described in Tipping (2001) as hierarcical priors is used, since it has desirable variable selection properties (Korobilis, 2013).

$$\mu_{\beta_j \to p(y_t|\beta)} = p\left(\beta_j\right) \prod_{s=1, s\neq t}^{T} \mu_{p(y_s|\beta)\to\beta_j} \tag{7}$$

We can estimate messages 6 and 7 above using the following iterative scheme, for $r = 1, \ldots, R$ where r denotes the order of iteration.

$$\mu^{(r+1)}_{p(y_t|\beta)\to\beta_i} = \int p\left(y_t|\beta\right) \prod_{j=1, j\neq i}^{p} \mu^{(r)}_{\beta_j\to p(y_t|\beta)} d\beta_{j\neq i} \tag{8}$$

$$\mu^{(r+1)}_{\beta_j \to p(y_t|\beta)} = p\left(\beta_j\right) \prod_{s=1, s\neq t}^{T} \mu^{(r)}_{p(y_s|\beta)\to\beta_j} \tag{9}$$

The GAMP algorithm employs Gaussian and Taylor series approximations which are based on asymptotic results such that as the number of parameters to be estimated increases, the analytical solutions derived from above iterations for the first two moments of the predictors become more reliable.[7] As in the case of the coefficient estimation of the exogenous predictors, stochastic volatility (σ_t^2) estimation is accomplished through data-driven priors without resorting to any form of Markov based dependence to σ_{t-1}^2.

3.2.2. Bayesian model averaging (TVP-BMA)

We employ time-varying parameter Bayesian model averaging approach of Groen et al., (2013) which incorporates model uncertainty as the relationship between housing returns and predictor variables is likely to have changed over time (Wei and Cao, 2017).

In particular, the TVP-BMA specification takes the following form:

$$\begin{aligned} y_{t+h} &= \sum_{j=1}^{p} x_{jt} s_j \beta_{jt} + \varepsilon_{t+h} \\ \beta_t &= \beta_{t-1} + \eta_t \end{aligned} \tag{10}$$

where β_{jt}'s are time-varying regression parameters and s_j is an indicator variable such that when $s_j = 0$ then j^{th} explanatory variable is eliminated from the regression in all periods, while $s_j = 1$ the predictor is included in the model. Since the full Bernoulli posterior of each parameter s_j is a sequence of zero and one values, the posterior mean can be interpreted as a well-defined probability of inclusion in the regression model of each variable j. Hence, this probability can be used for variable selection.[8]

[7] For an illustration of a simplified version of GAMP algorithm to derive mean and variance of β , see Korobilis (2019).

[8] It is assumed that the probabilities s_j have a Bernoulli prior with prior inclusion probability of each variable equal to 0.5.

3.2.3. *Ridge regression (RIDGE)*

The RIDGE regression implements a form of shrinkage by adding a constraint on the size of the coefficients to the usual sum of the squares minimization problem. As proposed by Hoerl and Kennard (1970), the RIDGE estimator is especially good at improving the least-squares estimate when multi-collinearity is present. Hence, it reduces the estimation variance by tilting the estimated parameters towards zero. Specifically, the RIDGE coefficients are obtained by solving the following problem:

$$\hat{\beta}^{ridge} = \min_{\beta} \quad ||Y - X\beta|| + \lambda \sum_{i=1}^{M} \beta_i^2, \tag{11}$$

where β is a M dimensional vector and $||Y - X\beta||$ shows ℓ_2-norm penalty. The parameter λ controls the degree of shrinkage; that is, the higher λ the closer to zero are the β_i, but they are never exactly zero.[9]

3.2.4. *Least absolute shrinkage operator (LASSO)*

We also employ the LASSO, which was proposed by Tibshirani (1996), and can be represented as a penalized regression problem. However, LASSO imposes an ℓ_1-norm penalty on the regression coefficients, rather than an ℓ_2-norm penalty in contrast to the ridge estimator. This penalty results in (possible) shrinkage of coefficients (called $\hat{\beta}^{lasso}$ below) to zero. The LASSO estimator is given below:

$$\hat{\beta}^{lasso} = \min_{\beta} \quad ||Y - X\beta||_2 + \lambda \sum_{j=1}^{N} |\beta_j|, \tag{12}$$

where λ is a tuning parameter that governs the strength of the ℓ_1-norm penalty. Since the objective function in the LASSO is not differentiable, numerical optimization techniques must be implemented when estimating $\hat{\beta}^{lasso}$. [10] However, one of the limitations of the LASSO approach is that the number of selected variables is bounded by the sample size. For example, if N > T, the LASSO yields at most N non-zero coefficients[11]. The variables associated with these non-zero coefficients constitute our set of predictors in our forecasting experiment.

3.2.5. *Elastic net (ENET)*

The LASSO is naturally ideal for situations where the "true" model includes several zero coefficients. However, Tibshirani (1996) reveals that the LASSO predictive performance is often weaker than those constructed by ridge regression in the presence of highly-correlated predictors. Zou and Hastie (2005)

[9] We perform 5-fold cross-validation over different values of λ and select the largest value of λ such that the mean squared error is achieved its minimum.

[10] For instance, we utilize an efficient iterative algorithm called the "shooting algorithm" which is introduced by Fu (1998).

[11] See Swanson (2016) for further discussion.

overcome this issue by introducing a hybrid form of the LASSO and ridge estimators, called the elastic net (ENET) estimator. The ENET estimator is defined as follows:

$$\hat{\beta}^{EN} = \min_{\beta} \quad ||Y - X\beta||_2 + \lambda_1 \sum_{j=1}^{N} |\beta_j| + \lambda_2 \sum_{j=1}^{N} {\beta_j}^2, \tag{13}$$

where there are now two tuning parameters, λ_1 and λ_2 controlling the two penalty functions. In addition, the ENET estimator results in possible shrinkage of coefficients to zero, although in cases where N > T, the ENET can produce more than N non-zero coefficients.

3.3. Forecasting experiments

We evaluate to forecasting performance of the sentiment indexes by using a recursive forecasting scheme, expanding the estimation window prior to the construction of each new forecast.[12] We run predictive regressions of the type commonly used in the forecasting literature, formulated as

$$y_{t+1} = \alpha + \beta S_t^j + \phi Z_t + u_{t+1} \tag{14}$$

where y_t represents the housing return and S_t^j is alternatively PCA- and PLS-based sentiment indexes. Z_t includes IP, CPI, IR, and SMR to take into account most of the relevant information about future house price returns contained in economic fundamentals. Finally, u_{t+1} represents error term. We reserve the period 2006:01-2010:12 to initial estimation period, and out-of-sample forecasts are computed over the period 2011:01 - 2018:12. For each month, we produce a sequence of six h-month-ahead forecasts, i.e., h=1, 2, 3, 6, 9, 12. Furthermore, we implement the equality of mean squared forecast error (MSFE) test of Harvey et al., (1997) to evaluate the forecast performance of the proposed models relative to our benchmark random-walk (RW) model.

As discussed at length by Bai and Ng (2008), Kuzin et al., (2011, 2013), Kim and Swanson (2014, 2018), Cepni and Guney (2019a), and Cepni et al., (2019b, 2020b), it is important to choose appropriate predictors prior to estimation of predictive regressions. The reason is that model and parameter uncertainty may adversely impact the marginal predictive content of explanatory variables. For this reason, we implement alternative time-varying parameter shrinkage models as discussed in Section 3.2. Accordingly, for each month, we choose indicators from the set of variables that includes IP, CPI, IR, SMR, and PLS-based sentiment index. Similarly, we implement a forecasting exercise that chooses from IP, CPI, IR, SMR, and the PCA-based sentiment index and compare the forecast performance of the models that contain the information from the two alternative investor sentiment indexes.

[12]This mitigates any concern over possible look-ahead bias.

4. Results

4.1. Main findings

Figure 1 plots the PCA- and PLS-based investor sentiment indexes. Both indexes stay high around the years 2007 and 2015, which is consistent with high house price appreciation in China during these periods. On the other hand, the PLS-based sentiment index show consistently higher values but lower volatility than the PCA-based sentiment index between the years 2010 and 2014. The reason for stable sentiment during these years might be that the central government mandated the local governments of major cities to introduce housing purchasing restrictions because of the fear of the housing bubble (Wu (2015)).

– Insert Figure 1 about here. –

Table 1 reports the out-of-sample forecasting results based on alternative model specifications. The results show that the MSFEs of the models that include macroeconomic variables and sentiment indexes generally increase with the forecast horizon. This finding shows that sentiment captures significant predictive information, particularly for shorter forecast horizons. Also, virtually most of the entries are higher than unity, implying that alternative specifications based on standard OLS estimation do not produce better forecasts than the benchmark RW model especially at longer forecast horizons (h= 6, 9, 12). However, we find that the model that includes the CPI, IP, IR, SMR, and the Investor.Sent.PLS index always provides the lowest MSFEs compared to the alternative model specification that comprises the PCA-based sentiment index. Hence, the PLS- based index contains more relevant information for the predictability of housing returns than the PCA-based sentiment index.[13]

Furthermore, when we let the time-varying parameter models choose from all key control variables including CPI, IP, IR, SMR, and the Investor.Sent.PLS index, the results in Table 1 are very encouraging for the use of time-varying parameter models that allow for model selection and parameter shifts.[14] In particular, TVP-BMA and TVP-GAMP models seem to be improving a lot over the benchmark RW model and the models that include economic fundamentals. This observation seems to suggest that TVP-BMA and TVP-GAMP models are flexible enough to capture both structural change and utilize the information in a set of predictors simultaneously as suggested by Korobilis (2019). While the predictive performance of the

[13]Note that, we also developed a sentiment index using the reduced-rank approach, originally developed by Anderson (1951), with Reinsel and Velu (1998) providing a book-level analysis on its properties and applications, and Huang et al., (2019) applying it financial data more recently. Mathematically, the RRA shrinks the dimension of factor space by imposing a rank restriction on regression coefficients, to reduce a large number of regressors to a small number of linear combinations. While the RRA-based sentiment index is found to outperform the PCA-based index beyond h=1, the former is always outperformed by sentiment index derived using the PLS for all the forecasting horizons considered. Complete details of these results are available upon request from the authors.

[14]Moreover, the results in Table A1 in the Appendix show that the superiority of Investor.Sent.PLS index continues to hold under the shrinkage models as well, since when we choose indicators from the CPI, IP, IR, SMR, and the Investor.Sent.PLS index, the MSFEs produced are lower compared to the corresponding MSFEs under the shrinkage approaches which select variables from CPI, IP, IR, SMR, and the Investor.Sent.PCA index.

TVP-BMA model is particularly notable for shorter forecast horizons (h= 1, 2, 3), the TVP-GAMP model always provides the lowest MSFEs and attains the top rank at relatively longer forecast horizons (h= 3, 6, 9). This observation is further supported by the predictive accuracy test of Harvey et al., (1997), which in turn implies statistically significant improvements in forecast accuracy compared to the RW model.

– Insert Table 1 about here. –

We also compare the predictive performance of time-varying model specifications that select indicators from the set of variables that includes CPI, IP, IR, SMR, and all the six investor sentiment proxies. As reported in Table 2, TVP-BMA and TVP-GAMP models retain their superiority in terms of out-of-sample forecasting, with the only exception of housing return forecasts from the LASSO regression at forecast horizon h=6. Put differently, although it is hard to pin down which variables contain useful information for predictions of housing returns when a large set of predictors is utilized, TVP-BMA and TVP-GAMP model are still good performing models over the alternative model specifications. On the other hand, the MSFE values of the best models in Table 2 are higher than those of the best models in Table 1, except for forecast period h=12. This observation suggests that optimally exploiting the six sentiment proxies based on the PLS procedure further improves the forecast performance of the competing models.

– Insert Table 2 about here. –

4.2. Robustness check - Regional segment evidence

It is well known that the housing sector in China is considered to be segmented (see for example, Hong and Li (2019b), Turner and Wessel (2019), Tsai and Chiang (2019)). Furthermore, with a focus on segmentation of the housing market, Goodman and Thibodeau (2007) suggest that an adequate understanding of the segmentation of the housing market would possibly improve the predictive accuracy of the models for forecasting the house price returns. The reason is that the real estate market is more prone to attract capital investment in areas with higher economic growth, rendering it much more vulnerable to speculation and investor sentiment. Hence, we further examine the relation between sentiment index and housing returns in different tier cities with different levels of economic developments. In their empirical study, Hong and Li (2019b) construct housing price indices for three tiers based on 70 large- and medium-size Chinese cities. Following their approach, we repeat our forecasting exercise on housing returns of the three tiers of cities.[15]

Table 3 shows that TVP-GAMP model is performing quite well, as it attains the top rank in 9 cases out of 18,[16], with this observation indicated via bold entries which correspond to the lowest MSFEs. Indeed, the

[15]Hong and Li (2019b) split the 70 major and medium-sized towns into three groups based on the official state declaration of the 'Town Classification Criteria.' The first-tier cities include Beijing, Shanghai, Guangzhou, and Shenzhen; the second-tier cities cover all the provincial capital cities and municipalities with independent planning status under the National Social and Economic Development Plan; and the other cities are classified into the third category. We thank Dr. Yun Hong and Dr. Yi Li for sharing their data in this regard.

[16]Recall that there are six forecast horizons and three tiers, meaning that we have a total of 18 specifications for each model.

MSFEs of the TVP-GAMP model are up to 28% lower than those associated with the RW model, especially at longer horizons. Similarly, the TVP-BMA keeps its superiority, yielding MSFE-best predictions in all tiers when considering only shorter forecast horizons (h=1, 2, 3). In particular, TVP-BMA results in 55% forecast improvement for tier 1 cities at h=1. These results underscore the importance of the specification of time variation in regression parameters and removing irrelevant variables when constructing predictions.

Furthermore, the results in Table 3 show that the model based on Investor.Sent.PLS outperforms the model based on Investor.Sent.PCA in a consistent fashion at all horizons and for all three tiers. On the other hand, the results in Panel C of Table 3 suggests that the inclusion of sentiment yields lower forecast errors of future housing returns for third-tier cities compared to the model which includes only economic fundamentals. This result is also consistent with Shiller (2007), who suggests that housing bubbles cannot be explained by economic fundamentals. Hence, a potential irrational component embedded in the sentiment index is needed to understand house price return dynamics.

– Insert Table 3 about here. –

In the Appendix, we also report the MSFEs of the forecasting exercise for all three tiers where time-varying parameter models are allowed to choose indicators from the set of variables that include CPI, IP, IR, SMR, and all the six investor sentiment proxies. The results in Table A2 in the Appendix show that TVP-GAMP and TVP-BMA models continue to perform well in all three-tier cities, implying that the importance of the individual predictors changes over time .

5. Conclusion

The financialization process of the housing market in China necessitates a more integrative approach, which incorporates swings in investor sentiment in the pricing mechanism. This study attempts to predict housing returns in China using aggregated and disaggregated proxies of investor sentiment in addition to macroeconomic fundamentals, based on various machine learning methods. Our findings suggest that investor sentiment has predictive power for housing returns primarily at short forecast horizons, and an aligned form of investor sentiment index obtained through PLS by combining equity market related individual proxies based on their relevance to housing market prices, is particularly useful in this regard compared to a PCA-based sentiment index. Moreover, the precision of the forecasts is further enhanced at all horizons by employing machine learning methods, where time variation in parameters and variable selection is allowed, thus underscoring the importance of the dynamic nature of the relationship between house prices and its various predictors: macroeconomic and behavioral.

The findings in this paper have several implications for practitioners and policy-makers. First, given that housing sector is one of the major pillars of the Chinese economy, accurate forecasting of housing returns has valuable implications to many stakeholders in the housing sector since it is heavily connected with both industries (such as home building, building materials, etc.) and the banking sector (including mortgage lending, home insurance, etc.). Second, the demonstrated predictability of the aligned investor sentiment index for housing returns in China corroborates the findings of Case and Shiller (1989, 1990), who conclude that housing markets are not fully efficient. Third, accurate forecasting of housing returns provides a near-term indicator of the health of the housing market for policy-makers to develop timely regulations in case of price anomalies, given that housing bubbles could turn into a bust with a potential contagion across financial sectors, and devastating impact on the macro-economy as witnessed during the recent global financial crisis of 2007-2008.

As part of future research, it would be interesting to extend our analysis of forecasting housing returns using investor sentiment and machine learning models to other developed and emerging economies. Moreover, further research should entail developing a sentiment index that relates specifically to housing-related decisions, and using this index, in turn, to forecast housing returns, as recently done for the US by Bork et al., (2020).

References

Aalbers, M. (2016). The financialization of housing: a political economy approach. London: Routledge Taylor & Francis Group.

Anderson, T. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. Annals of Mathematical Statistics, 22, 327–351.

Angelino, E., Johnson, M. J., Adams, R. P. (2016). Patterns of scalable Bayesian inference. Foundations and Trends in Machine Learning, 9(2-3), 119-247.

Bai, J., and Ng, S. (2008). Forecasting economic time series using targeted predictors. Journal of Econometrics, 146(2), 304-317.

Baker, M., and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. The journal of Finance, 61(4), 1645-1680.

Boivin, J., and Ng, S. (2006). Are more data always better for factor analysis? Journal of Econometrics, 132, 169-194.

Bork, L., and Møller, S.V. (2015). Forecasting house prices in the 50 states using Dynamic Model Averaging and Dynamic Model Selection. International Journal of Forecasting, 31(1), 63-78.

Bork, L., and Møller, S.V. (2018). Housing price forecastability: A factor analysis. Real Estate Economics, 46(3), 582-611.

Bork, L., , Møller, S.V., and Pedersen, T.Q. (2020). A New Index of Housing Sentiment. Management Science, 66(4), 1563–1583.

Case, K. E., and Shiller, R.J. (1989). The efficiency of the market for single family homes. American Economic Review, 79(1), 125–137.

Case, K. E., and Shiller, R.J. (1990). Forecasting prices and excess returns in the housing market. Real Estate Economics, 18(3), 253–273.

Cepni, O., and Guney, I.E. (2019a). Nowcasting emerging market's GDP: the importance of dimension reduction techniques. Applied Economics Letters, 26(20), 1670-1674.

Cepni, O., Guney, I.E., and Swanson, N.R. (2019b). Nowcasting and forecasting GDP in emerging markets using global financial and macroeconomic diffusion indexes. International Journal of Forecasting, 35(2), 555-572.

Cepni, O., Guney, I.E., Gupta, R., and Wohar, M.E. (2020a). The role of an aligned investor sentiment index in predicting bond risk premia of the US. Journal of Financial Markets, 100541.

Cepni, O., Guney, I.E., and Swanson, N.R. (2020b). Forecasting and nowcasting emerging market GDP growth rates: The role of latent global economic policy uncertainty and macroeconomic data surprise factors. Journal of Forecasting, 39(1), 18-36.

Chow S.-C., Cunado, J., Gupta, R., and Wong W.-K. (2018). Causal relationships between economic policy uncertainty and housing market returns in China and India: evidence from linear and nonlinear panel and time series models. Studies in Nonlinear Dynamics & Econometrics, 22(2), 1-15.

Deng, L., and Chen, J. (2019). Market development, state intervention, and the dynamics of new housing investment in China. Journal of Urban Affairs, 41(2), 223–47.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). The elements of statistical learning (Vol. 1, No. 10). New York: Springer series in statistics.

Fu, W.J. (1998). Penalized regressions: the bridge versus the lasso. Journal of computational and graphical statistics, 7(3), 397-416.

Goodman, A.C. and Thibodeau, T.G. (2007). The Spatial Proximity of Metropolitan Area Housing Submarkets. Real Estate Economics 35(2), 209–32.

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. International Journal of forecasting, 13(2), 281-291.

Hoerl, A.E., and Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67.

Hong, Y., and Li, Y. (2019a). House price and the stock market prices dynamics: evidence from China using a wavelet approach. Applied Economics Letters. DOI: https://doi.org/10.1080/13504851.2019.1649359.

Hong, Y., and Li, Y. (2019b). Housing prices and investor sentiment dynamics: Evidence from China using a wavelet approach. Finance Research Letters. DOI: https://doi.org/10.1016/j.frl.2019.09.015.

Hsing, Y.-T. (2010). The great urban transformation: Politics of land and property in China. Oxford, UK: Oxford University Press.

Huang, D., and Li, J., and Zhou, G. (2019). Shrinking Factor Dimension: A Reduced-Rank Approach. Available at SSRN: https://ssrn.com/abstract=3205697.

Kelly, B., and Pruitt, S. (2013). Market expectations in the cross-section of present values. Journal of Finance, 68, 1721–1756.

Kelly, B., and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. Journal of Econometrics, 186(2), 294-316.

Kim, H.H., and Swanson, N.R. (2014). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. Journal of Econometrics, 178, 352-367.

Kim, H.H., and Swanson, N.R. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. International Journal of Forecasting, 34(2), 339-354.

Korobilis, D. (2013). Hierarchical shrinkage priors for dynamic regressions with many predictors. International Journal of Forecasting, 29, 43-59.

Korobilis, D. (2019). High-dimensional macroeconomic forecasting using message passing algorithms. Journal of Business and Economic Statistics, 1-12.

Kuzin, V., Marcellino, M., and Schumacher, C. (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. International Journal of Forecasting, 27(2), 529-542.

Kuzin, V., Marcellino, M., and Schumacher, C. (2013). Pooling versus model selection for nowcasting GDP with many predictors: Empirical evidence for six industrialized countries. Journal of Applied Econometrics, 28(3), 392-411.

Lin, G.C.S. (2014). China's landed urbanization: Neoliberalizing politics, land commodification, and municipal finance in the growth of metropolises. Environment and Planning A: Economy and Space, 46(8), 1814–35.

Logan, J.R., Bian, Y.J., and Bian, F.Q. (1999). Housing inequality in urban China in the 1990s. International Journal of Urban and Regional Research, 23(1), 7–25.

Pearl, J. (1982). Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach. Proceedings of the National Conference on Artificial Intelligence, Pittsburgh, PA, USA, 133–136.

Reinsel, G., and Velu, R. (1998). Multivariate reduced rank regression. Lecture Notes in Statistics. New York: Springer.

Salisu, A.A., and Gupta, R. (Forthcoming). How do Housing Returns in Emerging Countries Respond to Oil Shocks? A MIDAS Touch. Emerging Markets Finance and Trade.

Shiller, R.J. (2007). Understanding recent trends in house prices and home ownership. Proceedings - Economic Policy Symposium - Jackson Hole, Federal Reserve Bank of Kansas City, 89-123.

Su, C.-W., Cai, X.-Y., and Tao, R. (2020). Can Stock Investor Sentiment Be Contagious in China? Sustainability, 12, 1571.

Swanson, N.R. (2016). Comment on: In sample inference and forecasting in misspecified factor models. Journal of Business and Economic Statistics, 34, 348-353.

Tang, K., and Xiong, W. (2012). Index investment and the financialization of commodities. Financial Analysts Journal, 68(6), 54–74.

Tao, R., Su, F.B., Liu, M.X., and Cao, G.Z. (2010). Land leasing and local public finance in China's regional development: Evidence from prefecture-level cities. Urban Studies, 47(10), 2217–2236.

Theurillat, T., Lenzer, Jr. J.H., and Zhan, H. (2016). The Increasing Financialization of China's Urbanization. Issues Studies, 52(04), 1640002.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

Tipping, M.E. (2001). Sparse Bayesian learning and the relevance vector machine. Journal of Machine Learning Research, 1, 211-244.

Tsai, I.C., and Chiang, S.H. (2019). Exuberance and spillovers in housing markets: Evidence from first-and second-tier cities in China. Regional Science and Urban Economics, 77, 75-86.

Turner, L.M., and Wessel, T. (2019). Housing market filtering in the Oslo region: pro-market housing policies in a Nordic welfare-state context. International Journal of Housing Policy, 19(4), 483-508.

Wei, Y., and Cao, Y. (2017). Forecasting house prices using dynamic model averaging approach: Evidence from China. Economic Modelling, 61, 147–155.

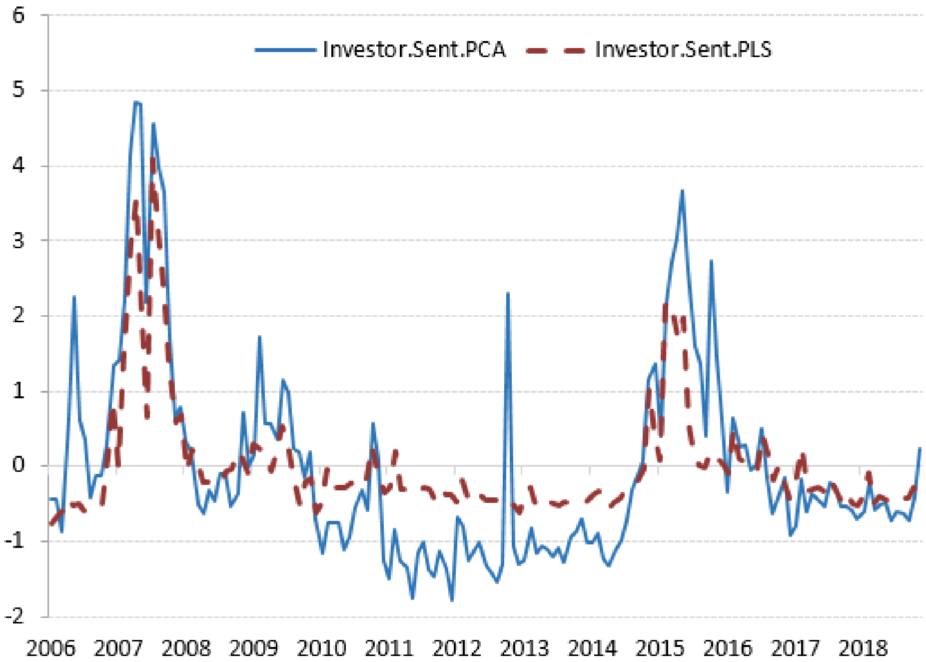
Wold, H., (1966). Estimation of principal component and related models by iterative least squares. In P. R. Krishnaiah (Ed.), Multivariate analysis, New York: Academic Press, 391–420.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, and V. Cappecchi (eds.), Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling. New York: Academic Press.

Wu, F. (2015). Commodification and housing market cycles in Chinese cities. International Journal of Housing Policy, 15(1), 6-26.

Wu, F., Chen, J., Pan, F., Gallent, N., and Zhang, F. (2020). Assetization: The Chinese Path to Housing Financialization. Annals of the American Association of Geographers. DOI: 10.1080/24694452.2020.1715195.

Figure 1: Sentiment Indexes



Notes: The figure plots the sentiment indexes. While dashed line represents the PCA based sentiment index, the solid line shows the PLS-based sentiment index. We multiply the Investor.Sent.PLS index with ten in order to plot the sentiment indexes in a common scale.

Table 1: Out-of-sample forecasting of housing returns based on alternative model specifications

Specification Type	h=1	h=2	h=3	h=6	h=9	h=12
RW	0.251	0.257	0.264	0.280	0.290	0.287
RW+ CPI+IP+IR+SMR	0.976*	0.992*	0.993	1.027	1.023	1.048
RW+ CPI+IP+IR+SMR+Investor.Sent.PCA	0.977*	0.998	1.007	1.049	1.058	1.086
RW+ CPI+IP+IR+SMR+Investor.Sent.PLS	0.970*	0.992	0.997	1.031	1.038	1.060
ENET	0.968*	0.966*	0.941**	0.858***	0.820***	0.957**
RIDGE	0.962**	0.965**	0.960**	0.903***	0.875***	1.000
LASSO	0.965*	0.966*	0.941**	0.844***	0.812***	0.955**
TVP-BMA	**0.558*** ***	**0.818*** ***	**0.743*** ***	1.038	1.513	1.981
TVP-GAMP	0.946***	0.980*	0.954**	**0.826*** ***	**0.754*** ***	**0.923*** ***

Note: Entries in the first row of the table are point MSFEs based on the benchmark random walk (RW) model, while the rest are relative MSFEs. Hence, a value of less than unity indicates that a particular model and estimation method is more accurate than that based on the RW model, for a given forecast horizon. Entries superscripted with an asterisk (** = 5% level; * = 10% level) are significantly superior than the RW model, based on the predictive accuracy test of Harvey et al., (1997). Entries that are yield the smallest MSFE are shown in bold.

Table 2: Out-of-sample forecasting of housing returns based on alternative model specifications with six sentiment proxies

Specification Type	h=1	h=2	h=3	h=6	h=9	h=12
RW	0.251	0.257	0.264	0.280	0.290	0.287
ENET	0.968*	0.959**	0.941**	0.867***	0.819***	0.965**
RIDGE	0.965**	0.954**	0.933***	0.894***	0.853***	0.958**
LASSO	0.964**	0.953**	0.954**	**0.854*** ***	0.825***	0.959**
TVP-BMA	**0.794*** ***	**0.850*** ***	1.078	1.182	1.312	1.676
TVP-GAMP	1.003	0.886***	**0.927*** ***	0.915***	**0.810*** ***	**0.899*** ***

Note: Entries in the first row of the table are point MSFEs based on the benchmark random walk (RW) model, while the rest are relative MSFEs. Hence, a value of less than unity indicates that a particular model and estimation method is more accurate than that based on the RW model, for a given forecast horizon. Entries superscripted with an asterisk (** = 5% level; * = 10% level) are significantly superior than the RW model, based on the predictive accuracy test of Harvey et al., (1997). Entries that are yield the smallest MSFE are shown in bold.

Table 3: Out-of-sample forecasting of housing returns based on alternative model specifications for three tier cities

Specification Type	h=1	h=2	h=3	h=6	h=9	h=12
PANEL A: TIER 1						
RW	0.247	0.254	0.260	0.276	0.286	0.283
RW+ CPI+IP+IR+SMR	0.976*	0.992	0.993	1.027	1.023	1.048
RW+ CPI+IP+IR+SMR+Investor.Sent.PCA	0.977*	0.998	1.007	1.048	1.058	1.086
RW+ CPI+IP+IR+SMR+Investor.Sent.PLS	0.970*	0.992*	0.997	1.031	1.038	1.060
ENET	0.970*	0.962*	0.936**	0.864***	0.817***	0.967**
RIDGE	0.966*	0.963**	0.961**	0.890***	0.874***	0.991*
LASSO	0.971**	0.974**	0.946***	0.836***	0.812***	0.956**
TVP-BMA	**0.549*****	**0.794*****	**0.747******	2.357	1.418	1.589
TVP-GAMP	0.948***	0.978*	0.955***	**0.824*****	**0.750*****	**0.929*****
PANEL B: TIER 2						
RW	0.308	0.316	0.325	0.345	0.357	0.357
RW+ CPI+IP+IR+SMR	0.962*	0.978*	0.980*	1.006	1.006	1.022
RW+ CPI+IP+IR+SMR+Investor.Sent.PCA	0.973*	0.993	1.003	1.040	1.007	1.069
RW+ CPI+IP+IR+SMR+Investor.Sent.PLS	0.967**	0.987*	0.993	1.023	1.029	1.041
ENET	0.986*	0.964**	0.953**	0.878***	0.853***	0.968**
RIDGE	0.970*	0.967**	0.972**	0.908***	0.892***	0.993
LASSO	0.982*	0.959**	0.946**	0.868***	0.843***	0.959**
TVP-BMA	**0.585*****	**0.745*****	**0.901*****	1.193	1.505	1.701
TVP-GAMP	0.999	0.985*	0.941**	**0.856*****	**0.779*****	**0.947*****
PANEL C: TIER 3						
RW	0.198	0.203	0.208	0.219	0.227	0.222
RW+ CPI+IP+IR+SMR	1.000	1.012	1.013	1.057	1.049	1.086
RW+ CPI+IP+IR+SMR+Investor.Sent.PCA	0.981*	1.002	1.010	1.057	1.062	1.105
RW+ CPI+IP+IR+SMR+Investor.Sent.PLS	0.968*	0.990*	0.995	1.033	1.039	1.075
ENET	0.978*	0.999*	0.970*	0.848***	0.779***	0.964**
RIDGE	0.976**	1.002	0.996	0.904***	0.850***	0.991
LASSO	0.974**	1.005	0.968*	0.831***	0.780***	0.962**
TVP-BMA	**0.567*****	**0.857*****	**0.800*****	1.112	1.261	1.264
TVP-GAMP	0.905**	0.958**	0.933**	**0.795*****	**0.728*****	**0.915*****

Note: Entries in the first row of the table are point MSFEs based on the benchmark random walk (RW) model, while the rest are relative MSFEs. Hence, a value of less than unity indicates that a particular model and estimation method is more accurate than that based on the RW model, for a given forecast horizon. Entries superscripted with an asterisk (** = 5% level; * = 10% level) are significantly superior than the RW model, based on the predictive accuracy test of Harvey et al., (1997). Entries that are yield the smallest MSFE are shown in bold.

Appendix

Table A1: MSFEs results of shrinkage models with alternative investor sentiment indexes

Specification Type	h=1	h=2	h=3	h=6	h=9	h=12
RW	0.246	0.251	0.268	0.286	0.301	0.307
ENET - PCA	0.968	0.971	0.961	1.012	1.079	1.028
RIDGE - PCA	0.961	0.961	0.958	1.036	1.095	1.049
LASSO - PCA	0.961	0.944	0.973	1.011	1.081	1.039
TVP-BMA - PCA	0.577	0.826	0.800	1.080	1.418	1.619
TVP-GAMP - PCA	0.921	0.957	0.928	0.848	0.777	0.955
ENET - PLS	0.968	0.966	0.941	0.858	0.820	0.957
RIDGE - PLS	0.962	0.965	0.960	0.903	0.875	1.000
LASSO - PLS	0.965	0.966	0.941	0.844	0.812	0.955
TVP-BMA - PLS	**0.558**	**0.818**	**0.743**	1.038	1.513	1.981
TVP-GAMP - PLS	0.946	0.980	0.954	**0.826**	**0.754**	**0.923**

Note: Entries in the first row of the table are point MSFEs based on the benchmark random walk (RW) model, while the rest are relative MSFEs. Hence, a value of less than unity indicates that a particular model and estimation method is more accurate than that based on the RW model, for a given forecast horizon. "-PLS" extended models show that the corresponding shrinkage model chooses indicators from the set of variables that includes IP, CPI, IR, SMR, and PLS-based sentiment index. Similarly, "-PCA" extended models indicate that the corresponding shrinkage model selects variables from IP, CPI, IR, SMR, and the PCA-based sentiment index. Entries that are yield the smallest MSFE are shown in bold.

Table A2: Out-of-sample forecasting of housing returns based on alternative model specifications with six individual proxies for three tier cities

Specification Type	h=1	h=2	h=3	h=6	h=9	h=12
PANEL A: TIER 1						
RW	0.247	0.254	0.260	0.276	0.286	0.283
ENET	0.966**	0.950**	0.934**	0.875***	0.818***	0.963**
RIDGE	0.973*	0.951**	0.928***	0.911***	0.848***	0.965**
LASSO	0.966**	0.951**	0.940***	**0.855*****	0.837***	0.959**
TVP-BMA	**0.771*****	**0.824*****	0.982*	1.259	1.159	1.682
TVP-GAMP	1.002	0.881***	**0.920*****	0.907***	**0.811*****	**0.904*****
PANEL B: TIER 2						
RW	0.308	0.316	0.325	0.345	0.357	0.357
ENET	0.971*	0.963**	0.954**	0.888***	0.852***	0.973**
RIDGE	0.970*	0.957**	0.948**	0.909***	0.875***	0.969**
LASSO	0.977*	0.956**	0.955**	**0.880*****	0.869***	0.969**
TVP-BMA	**0.634*****	**0.861*****	1.034	1.319	1.291	1.784
TVP-GAMP	0.982*	0.963**	**0.918*****	0.912***	**0.843*****	**0.924*****
PANEL C: TIER 3						
RW	0.198	0.203	0.208	0.219	0.227	0.222
ENET	0.987*	0.957**	0.942**	0.857***	0.811***	0.949**
RIDGE	1.000	0.965**	0.934**	0.884***	**0.805*****	0.948**
LASSO	0.997	0.963**	0.957**	0.852***	0.812***	0.938***
TVP-BMA	**0.766*****	**0.925*****	1.030	1.203	1.342	1.541
TVP-GAMP	1.033	0.836***	**0.901*****	**0.835*****	0.806***	**0.895*****

Note: Entries in the first row of the table are point MSFEs based on the benchmark random walk (RW) model, while the rest are relative MSFEs. Hence, a value of less than unity indicates that a particular model and estimation method is more accurate than that based on the RW model, for a given forecast horizon. Entries superscripted with an asterisk (** = 5% level; * = 10% level) are significantly superior than the RW model, based on the predictive accuracy test of Harvey et al., (1997). Entries that are yield the smallest MSFE are shown in bold.