

The Impact of Government Policies on the Sentiment of Public Opinion in the Real Estate Sector

Zitong Gao^{1*}, Qiaoyi Gai²

¹*Nanjing University, Nanjing, China*

²*Qingdao Dewey Experimental School, Qingdao, China*

**Corresponding Author. Email: 211098357@smail.nju.edu.cn*

Abstract. Since the beginning of the epidemic, the real estate industry has gradually begun to decline, and the national housing climate index has also become a downward trend. At present, the real estate industry is one of the topics of much attention in China. And the country also constantly introducing real estate policies to reverse the current real estate adversity. Although many scholars have made a research on the relationship between market sentiment and urban housing prices, the field of academic papers lacks the exploration of the relationship between public sentiment and the National Real Estate Development Index. Therefore, this paper focuses on the relationship between the public sentiment index and the National Housing Prosperity Index.

Keywords: component, real estate, policies, event study, public sentiment

1. Introduction

1.1. Research background

Since the beginning of the epidemic, the real estate industry has gradually begun to decline, and the national housing climate index has also become a downward trend. Moreover, the housing prices in many places have fallen to different degrees. The author, Tai Yuhong, studied the influence of principal component analysis (PCA) on the impact of housing price fluctuations in different tier cities in China by constructing (a PCA) and vector autoregression (VAR) model [1]. It also shows that the influence of the real estate market sentiment on housing price fluctuation is different under different degrees of fluctuation in economic policy uncertainty. Among them, housing price fluctuations in first-tier cities are more sensitive to market sentiment, while those in second-tier cities are relatively stable.

1.2. Literature review

At present, the real estate industry is one of the topics of much attention in China. With the development of Internet social platforms, more and more people are willing to express their views and express their attitudes on the Internet. Therefore, at this stage, there are a large number of comments and opinions on today's real estate industry on the Internet [2]. Social media not only

provides rich information products but also changes the way information is acquired and released and expands the channels of information exchange in the capital market. The heterogeneity and free expression of social media content may have enhanced its information content and provided a new perspective and tool for capital market analysis.

At this stage, the country is also constantly introducing real estate policies to reverse the current real estate adversity. When a series of new policies are released, people will also release their views on the policy and their attitudes towards real estate after the policy release. At this point, the market sentiment will change [3]. Market sentiment affects people's decisions and then affects the behavior of the real estate market, and then has an impact on the housing price [4]. At the same time, the mood of media reports can affect public expectations and thus affect the fluctuation of housing prices [5]. In order to make the real estate industry recover and formulate more appropriate policies, many scholars have conducted research on the trend of real estate and related influencing factors. At the same time, many scholars have conducted research on the relationship between market sentiment and urban housing prices. The author Kuang Weida used data from 2000 to 2015 to study the impact of heterogeneous media sentiment on housing price fluctuations in 112 large and medium-sized cities in China using the two-stage least squares method (2 SLS) [6]. The study found that media sentiment has a lag effect and superposition effect and has a significant impact on the fluctuation of housing prices. In particular, the sentiment of opinion leaders in Weibo and the reports of the central media can form a strong opinion environment among investors, thus having a greater impact on the fluctuation of housing prices.

In the other fields, there is still some research about the relationship between the development of the work and the emotions. For example, in the stock industry, the author Jiang Fuwei uses text sentiment analysis technology to extract investor sentiment information from financial news and explore the ability of this emotional information to predict the stock market [7].

1.3. Research significance

However, so far, the field of academic papers has not explored the relationship between public sentiment and the National Housing Prosperity Index. Therefore, this paper focuses on the relationship between the public sentiment index and the National Housing Prosperity Index. The paper mainly uses data analysis to do the research and the event study to explore the topic in depth. Therefore, it could help the government make more proper policies. In this way, the real estate will be more likely to recover.

1.4. Related concepts and theories

1.4.1. National Housing Prosperity Index

First of all, the paper needs to explain some concepts and theories used in the passage. National Housing Prosperity Index can be known by exploring. Its preparation method is according to the economic cycle, fluctuation theory [8] and climate index principle, the synthetic index calculation method, using time series, multiple statistics, and econometrics analysis method, with real estate development investment as the benchmark index, selected the real estate investment, capital, area, sales and other relevant indicators, and eliminate the influence of seasonal factors, contains the random factors. The National Housing Prosperity Index is a comprehensive reflect the national real estate industry development of overall index, from the land, capital, development, and market demand shows the national real estate industry's basic operation status, and volatility, predicts the

future trend, provide early warning mechanism for national macroeconomic regulation and control, provide statistical information for investors to choose investment opportunities. The national housing climate index is a weighted average comprehensive index composed of the index index of advanced, synchronous and lagging nature. According to the calculation method of the synthetic index, on the basis of calculating the classified index. And because of the article requirement, the paper finds some related real estate policies in Q1 2024. Its content is mainly about allowing some cities to cancel or reduce housing restrictions, cancel the ordinary residential and non-ordinary residential classification standards, promote the reformation of the real estate development financing method and commercial housing presale system, speed up the construction of the new model of real estate, just need to meet basic needs of the houses of the working class and support urban and rural residents diversification to improve the housing demand.

1.4.2. Commercial residential housing

Commercial residential housing is defined as commercial housing built by developers and sold through the market to individuals or units for residential purposes. In China, commercial residential housing is a very important part of the real estate market, which meets the housing needs of the general public. The property right of commercial residential housing is generally 70 years, counting from the time the land use right is ceded. After the expiration of the property right, according to the law, if the period of residential construction land use right expires, it will be automatically renewed. With the ever-changing regulatory policies in the real estate market, different cities and regions have different restrictions on the eligibility of home buyers, such as purchase restriction policies and down payment ratios. Commercial, residential housing not only solves the housing problem of the residents but also is closely related to the planning and development of the city and the socio-economic situation, which is one of the important parts of China's real estate market.

1.4.3. Year-on-year increase in sales of commercial residential housing

Year-over-year growth in commercial and residential home sales is the ratio of growth between commercial and residential home sales in the current reporting period and commercial and residential home sales in the same reporting period of the previous year. This ratio can reflect whether the real estate market has improved or deteriorated in sales over time. This indicator can help us understand the heat and trend of the real estate market. Below is a brief introduction to year-on-year residential sales growth and how it is calculated:

$$year-on-year\ growth\ rate = \left(\frac{Sales\ for\ the\ period - Sales\ for\ the\ same\ period\ of\ the\ previous\ year}{Sales\ for\ the\ same\ period\ of\ the\ previous\ year} \right) \times 100\%$$

The year-on-year growth rate of commercial residential housing sales is an important macroeconomic indicator that helps governments, enterprises, and investors understand the dynamics of the real estate market so that they can make decisions accordingly. For example, if the growth rate continues to rise, it may indicate that the real estate market is picking up; on the contrary, if the growth rate declines or even goes negative, it may mean that the market is facing challenges.

In addition, this indicator can help analyze the cyclical characteristics of the real estate market, predict future development trends, and provide a reference basis for relevant policymaking.

2. Data

2.1. Data collection

2.1.1. Sentiment index

It is observed that the Bilibili platform and Xiaohongshu have a considerable number of videos and posts related to the issue of China's real estate market turmoil, and the playback, pop-up, and comment volumes are high, which is representative and universal as a sample for public opinion analysis. Therefore, the paper used Python crawler and sentiment analysis technology to calculate the sentiment index of each comment and pop-up by introducing SnowNLP library, i.e., positive sentiment ratio, neutral sentiment ratio and negative sentiment ratio. Here the paper used the positive sentiment ratio as the sentiment index, and conduct a public opinion analysis on the current situation of real estate in China based on the pop-up data of related videos on Bilibili and related posts on Xiaohongshu. In the end, the paper crawled a total of 45,701 Bilibili pop-ups and 25,049 Xiaohongshu comments, and analyzed their monthly sentiment from January 2022 to June 2024 after sorting them according to time.

Meanwhile, in order to explore the degree of prosperity of China's real estate market, the paper collected China's national housing prosperity index from the CSMAR database from 1988 to June 2024, and in order to correspond with the sentiment index, the paper selected the monthly national housing prosperity index from January 2022 to June 2024.

Since the logic and steps of crawling and performing sentiment analysis with the program for Bilibili pop-ups and Xiaohongshu comments have a large similarity, I will take the Bilibili platform as an example to elaborate next.

2.1.1.1. Crawler

1) Video and Post Selection

In order to make the results of public opinion analysis more credible and generalized, the paper selects keywords with greater relevance to the real estate market, such as "real estate", "interest rate", "down payment", "provident fund", "buying a house", to search for videos and posts, by searching for these keywords we selected a total of 91 Bilibili videos and 64 Xiaohongshu posts related to the current situation of China's real estate.

2) Define the Function `get_bilibili_danmu` to Crawl the Video Pop-ups on Bilibili

In the Headers setting, in order to make the server think that the request is from a browser rather than a crawler, we simulate a user agent; then send a GET request to get the video information and convert it to a Python object through JSON parsing, then extract the video's cid (the pop-up screen's identifier) from the JSON data, and then construct a link to get the pop-up screen and send the request again; in order to deal with the special characters in the pop-up data, we encode the obtained pop-up data and convert it to byte type; then we use BeautifulSoup to parse the pop-up data in XML format and find all the pop-up tags, and save them in `danmu_list` for subsequent processing. Then we created four initialized lists for storing the popup information and traversed all the popup tags through a loop structure to extract the information, and finally created a DataFrame to store the popup information. Figure 1 shows the first part of the crawler source code.

```
def get_bilibili_danmu(v_url, v_result_file):
    headers = {'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko)', }
    print('The video URL is:', v_url)
    r1 = requests.get(url='https://api.bilibili.com/x/player/pagelist?bvid='+bv, headers=headers)
    html1 = r1.json()
    cid = html1['data'][0]['cid'] # Get the cid of the video
    print('The video's cid is:', cid)
    danmu_url = 'http://comment.bilibili.com/{}.xml'.format(cid)
    print('The pop-up URL is:', danmu_url)
    r2 = requests.get(danmu_url, headers=headers)
    html2 = r2.text.encode('raw_unicode_escape') # coding format
    soup = BS(html2, 'xml')
    danmu_list = soup.find_all('d')
    print('Crawled a total of {} pop-ups'.format(len(danmu_list)))
    video_url_list = [] # Video URL
    danmu_url_list = [] # Pop-up URL
    time_list = [] # Pop-up time
    text_list = [] # Pop-up content
    for d in danmu_list:
        data_split = d['p'].split(',') # comma-delimited
        temp_time = time.localtime(data_split[4]) # Convert Time Format
        danmu_time = time.strftime("%Y-%m-%d %H:%M:%S", temp_time)
        video_url_list.append(v_url)
        danmu_url_list.append(danmu_url)
        time_list.append(danmu_time)
        text_list.append(d.text)
        print('{}:{}'.format(danmu_time, d.text))
    df = pd.DataFrame()
    df['Video URL'] = video_url_list
    df['Pop-up URL'] = danmu_url_list
    df['Pop-up time'] = time_list
    df['Pop-up content'] = text_list
```

Figure 1. Excerpts from the source code-1

In order to avoid writing the CSV header multiple times, we designed a processing logic: if the file exists, there is no need to write the field header; if the file does not exist (indicating that it is the first time a new file is created), there is a need to write the field header. On this basis, the DataFrame is written to the CSV file. Figure 2 shows the second part of the crawler source code.

```
if os.path.exists(v_result_file): # If the file exists, there is no need to write the field title
    header = None
else: # If the file does not exist, it means it is the first time a new file is created and the field title needs to be written
    header = ['Video URL', 'Pop-up URL', 'Pop-up time', 'Pop-up content']
    df.to_csv(v_result_file, encoding='utf_8_sig', mode='a+', index=False, header=header)
```

Figure 2. Excerpts from the source code-2

3) Main Module

First build a list based on the found video BV serial numbers, then traverse the list of BV numbers and execute the get_bilibili_danmu function for each BV number. Crawl a set of Bilibili video pop-up messages and then store these messages into a CSV file. To ensure that each execution is a new data write, before each execution, it checks if a CSV file with the same name exists, and deletes it if it does. Figure 3 shows the third part of the crawler source code.

```
print('The crawler program starts executing!')
csv_file = 'Real estate pop-ups.csv'
if os.path.exists(csv_file):
    print('{} already exists, start deleting files'.format(csv_file))
    os.remove(csv_file)
    print('{} Deleted files'.format(csv_file))
bv_list = ['BV1m94y1M7bc', 'BV18v411s7XJ', 'BV1D3411a7yo', 'BV16P4y1p7m5', 'BV154y1j7Nk', 'BV1454y1p788', 'BV1ba4y1m7ho', 'BV1r14y1F7yN',
for bv in bv_list:
    get_bilibili_danmu(v_url='https://www.bilibili.com/video/{}'.format(bv), v_result_file='Real estate pop-ups.csv.csv')
print('The crawler program is executed!')
```

Figure 3. Excerpts from source code-3

2.1.1.2. Sentiment determination - marking pop-up content with SnowNLP

1) Definition of Sentiment Index Function

SnowNLP library is a Python library for sentiment analysis of Chinese text, it is especially suitable for processing Chinese text data. Compared to other sentiment analysis tools, SnowNLP focuses more on Chinese text processing, providing Chinese word splitting, sentiment analysis and other functions, is a powerful tool for Chinese sentiment analysis.

Using this library for sentiment analysis requires the definition of a decision marking function. The function accepts a comment list v_cmt_list as input. First, some variables are initialized to store the sentiment score, the label classification results, and the counts of positive and negative comments. Then it traverses the list of comments and performs sentiment analysis on each

comment: first use SnowNLP library to calculate the sentiment score of each comment; then set the comment labels according to the sentiment score and update the counts of positive and negative comments (we set the value of the sentiment score less than 0.5 to be negative, more than 0.5 to be positive, and otherwise to be neutral); then add the sentiment score and labels to the corresponding lists and calculate the percentage of positive, neutral and negative comments. Finally, create a DataFrame, add the sentiment scores and labels to it and save the DataFrame to an Excel file. Figure 4 shows the first part of the sentiment analysis source code.

```
def sentiment_analyse(v_cmt_list):
    score_list = []
    tag_list = []
    pos_count = 0
    neg_count = 0
    mid_count = 0
    for comment in v_cmt_list:
        tag = ''
        sentiments_score = SnowNLP(comment).sentiments
        if sentiments_score < 0.5:
            tag = 'Negative'
            neg_count += 1
        elif sentiments_score > 0.5:
            tag = 'Positive'
            pos_count += 1
        else:
            tag = 'Neutral'
            mid_count += 1
        score_list.append(sentiments_score)
        tag_list.append(tag)
    df['sentiment score'] = score_list
    df['analysis'] = tag_list
    grp = df['analysis'].value_counts()
    print('Positive and negative review stats:')
    print(grp)
    df.to_excel('Sentiment Scoring Results.xlsx', index=None)
    print('Sentiment analysis results have been generated: Sentiment Scoring Results.xlsx')
```

Figure 4. Excerpts from source code-4

2) Main Module

First read the CSV file 'real estate popup.csv' obtained by crawler and store it in DataFrame object, then read the "popup content" column, convert the value of the column to a list and store it in the comment list v_cmt_list, and then convert each element in the comment list to a string type and call the previously defined function to perform sentiment analysis on the comment list v_cmt_list. Convert each element of the comment list to a string type, and then call the previously defined function to analyze the sentiment of the comment list v_cmt_list. Meanwhile, in order to facilitate high-frequency word statistics later, we use spaces to concatenate all comment strings to form a single string v_cmt_str. Figure 5 shows the second part of the sentiment analysis source code.

```
df = pd.read_excel('Sentiment Scoring Results.xlsx')
v_cmt_list = df['Pop-up content'].values.tolist()
print('length of v_cmt_list is:{}'.format(len(v_cmt_list)))
v_cmt_list = [str(i) for i in v_cmt_list]
v_cmt_str = ' '.join(str(i) for i in v_cmt_list)
sentiment_analyse(v_cmt_list=v_cmt_list)
```

Figure 5. Excerpts from source code-5

3) High-frequency Word Statistics

First use the comment string v_cmt_str obtained in the previous step to perform jieba segmentation (using the exact mode and turning off the new word discovery function), then the segmentation results are connected into a string with commas and stored in the variable r. Then set the deactivation words to filter out some meaningless words, and then use the keyword extraction function to extract the top 20 high-frequency words in the comments and get their weights, and then use the keyword extraction function to extract the top 20 high-frequency words in the comments. At the same time, in order to make the results more targeted and analyzable, we only keep the words of

nouns, verbs, adjectives, noun names, noun places and other words. Figure 6 shows the high-frequency word statistics source code.

```
s1 = jieba.cut(v_cmt_str, cut_all=False, HMM=False)
r='.'.join(s1)
jieba.analyse.set_stop_words("stopwords.txt")
keywords_top20 = jieba.analyse.extract_tags(r, withWeight=True, topK=20, allowPOS=('', 'n', 'v', 'a', 'nr', 'ns'))
print("top20 keywords and weights: ")
pprint(keywords_top20)
```

Figure 6. Excerpts from source code-6

2.1.1.3. Data validation and presentation

Figure 7 shows the monthly sentiment index from January 2022 to June 2024 that we have compiled.

2022M01	0.4061	2023M04	0.5363
2022M02	0.5149	2023M05	0.5496
2022M03	0.4795	2023M06	0.5404
2022M04	0.399	2023M07	0.4845
2022M05	0.4	2023M08	0.4326
2022M06	0.411	2023M09	0.523
2022M07	0.416	2023M10	0.4346
2022M08	0.4762	2023M11	0.4442
2022M09	0.4173	2023M12	0.4113
2022M10	0.3212	2024M01	0.5466
2022M11	0.3865	2024M02	0.5199
2022M12	0.3298	2024M03	0.5256
2023M01	0.3647	2024M04	0.471
2023M02	0.4168	2024M05	0.5423
2023M03	0.4784	2024M06	0.5031

Figure 7. Sentiment index

In order to prove the credibility and reliability of this data, we searched another dataset in addition to the National Housing Prosperity Index. Considering that the videos and posts in Bilibili and Xiaohongshu have higher playback and comment volume when the search terms are "buying a house", "mortgage" and "down payment", we believe that there is a higher probability that the public's emotions and opinions will focus on their real demand for buying houses and ultimately be projected onto the sales of commercial residential houses in the real estate market. Therefore, we believe that if the year-on-year growth in sales of commercial residential housing has a higher correlation with the sentiment index, then the sentiment index we calculated through programming has greater credibility and research significance.

1) Collection of Year-on-Year Growth in Sales of Commercial Residential Housing

We consulted the National Bureau of Statistics to obtain the year-on-year growth in sales of commercial residential housing from February 2022 to June 2024, which is vacant in January of each year due to issues with the methodology used to calculate this data in China. The data is presented in the Figure 8.

2022M01	NA	2023M04	11.8
2022M02	-22.1	2023M05	11.9
2022M03	-25.6	2023M06	3.7
2022M04	-32.2	2023M07	0.7
2022M05	-34.5	2023M08	-1.5
2022M06	-31.8	2023M09	-3.2
2022M07	-31.4	2023M10	-3.7
2022M08	-30.3	2023M11	-4.3
2022M09	-28.6	2023M12	-6
2022M10	-28.2	2024M01	NA
2022M11	-28.4	2024M02	-32.7
2022M12	-28.3	2024M03	-30.7
2023M01	NA	2024M04	-31.1
2023M02	3.5	2024M05	-30.5
2023M03	7.1	2024M06	-26.9

Figure 8. Year-on-year growth in sales of commercial residential housing from NBS

2) Correlation Test

Subsequently, we conducted a correlation test between the two on econometric software and found a relatively strong time-series correlation. The implication is that there is a relatively strong correlation of 74.47% between the public opinion sentiment index of the current period and the sales growth rate of the housing market in the previous period. Therefore, we believe that the sentiment index we calculated with the program has a degree of credibility and is valuable for research. Figure 9 shows the results of the correlation test between the current sentiment index and the year-on-year increase in sales of commercial residential housing in the previous period.

Correlation					
	PI	S(-1)			
PI	1.000000	0.744651			
S(-1)	0.744651	1.000000			

Figure 9. Correlation of PI and S; PI: sentiment index, S: year-on-year growth in sales of commercial residential housing index

2.1.2. National Housing Prosperity Index (NHPI)

In order to obtain the National Housing Prosperity Index, we searched the CSMAR database searching the real estate industry development data panel and found the National Housing Prosperity Index from 1986 to June 2024 for our country. Figure 10 shows the data we collected.

2022M01	0.9683	2023M04	0.9478
2022M02	0.9693	2023M05	0.9456
2022M03	0.9666	2023M06	0.9406
2022M04	0.9589	2023M07	0.9378
2022M05	0.956	2023M08	0.9356
2022M06	0.954	2023M09	0.9344
2022M07	0.9526	2023M10	0.934
2022M08	0.9507	2023M11	0.9342
2022M09	0.9486	2023M12	0.9336
2022M10	0.947	2024M01	0.924
2022M11	0.9442	2024M02	0.9213
2022M12	0.9435	2024M03	0.9207
2023M01	0.9445	2024M04	0.9202
2023M02	0.9467	2024M05	0.9201
2023M03	0.9471	2024M06	0.9211

Figure 10. National Housing Prosperity Index

2.2. Data analysis

In order to better analyze the data we obtained, we used metrics software to draw dot plots and line plots for each series of data.

First observe the sentiment index. Figure 11 shows the dot plots and the line plots for sentiment index.

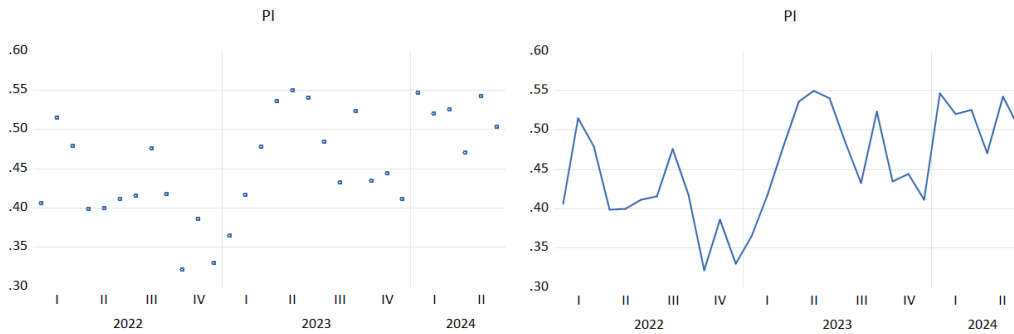


Figure 11. Dot plots and line plots for sentiment index

We can observe a strong similarity between the sentiment indexes in 2022 and 2023, characterized by a steady increase in the first and second quarters, a downward trend in the third quarter, a low in the fourth quarter, and then an upward trend in the first quarter of the following year. the difference between the indexes in 2022 and 2023 is that the average of the indexes is higher in 2023 than in 2022. In other words, there is a high likelihood that public opinion is more positive in 2023 than in 2022. However, we can observe that the trend of the sentiment index in 2024 is not quite in line with the previous two years, with a downward trend already in the first quarter and a brief rise in the second quarter in the month of May, followed by a rapid fall.

Next observe the National Housing Prosperity Index. Figure 12 shows the dot plots and the line plots for National Housing Prosperity Index.

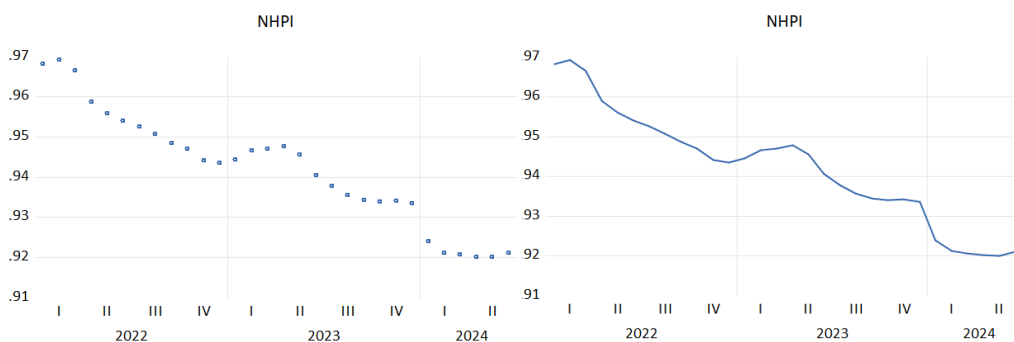


Figure 12. Dot plots and line plots for the National Housing Prosperity Index

The State Housing Prosperity Index has been generally trending downward for the past two and a half years. 2022 and 2023 both show an upward trend in the first quarter, but 2023 shows an upward trend in the fourth quarter, unlike the previous year. In 2024, weakness was evident in the first quarter, and it was not until the second quarter that a slight upward trend was seen.

It can be seen that in 2024 both series of data showed a different development trend from the previous one, so we believe that the real estate policies introduced by the country may have had a

certain degree of impact on the real estate industry in the country. For this reason, we decided to build an event study model to better study this phenomenon.

3. Event study model

3.1. Research event identification

In the first quarter of 2024, the State released a series of new real estate policies, the contents of which are mainly to allow some cities to abolish or reduce housing restrictions, abolish the classification standards for ordinary and non-ordinary residences, promote the reform of real estate development financing methods and the pre-sale system of commercial properties, accelerate the construction of new real estate models, satisfy the working class's immediate need for housing and support the diversified and improved housing needs of urban and rural residents.

There is a high probability that these real estate policies were introduced with the intention of increasing consumers' willingness to buy homes and improving the declining real estate market. But will these policies improve the mood of public opinion and the National Housing Prosperity Index? If they do, will their impact be positive or negative?

3.2. Model setting

We thus set the first quarter of 2024 as the time window and set January 2022 through December 2023 as the study sample for the benchmark model to fit the normal returns for the second quarter of 2024. The specific settings of the model are given below.

3.2.1. Benchmark model

3.2.1.1. Normal revenue projections

1) Time-series Test

In order to explore the impact of real estate policies on public opinion sentiment, we construct a dynamic panel model to capture the changes in the popular sentiment index. Specifically, we assume that the current period's popular sentiment index is affected not only by the previous period's popular sentiment index but also by the previous period's national housing sentiment index.

The sentiment index is denoted by PI, and the National Housing Prosperity Index is denoted by NHPI. The sample data are first tested to see if they are time series, which is done by performing the Granger causality test. The specific results of the Granger causality test are shown in Table 1:

Table 1. Results of the Granger causality test

Null Hypothesis(Lag:3)	Obs	F-statistic	Prob.
NHPI does not Granger Cause PI	27	1.44945	0.2583
PI does not Granger Cause NHPI		1.44187	0.2604

As a result, the hypothesis that both variables are not time series is rejected, proving that both are time series, and we determined the optimal number of lags for the PI to be the first period and the optimal number of lags for the NHPI to be the second period by performing a Correlogram test. Figures 13 and 14 show the time series tests for the sentiment index and the national housing prosperity index, respectively.

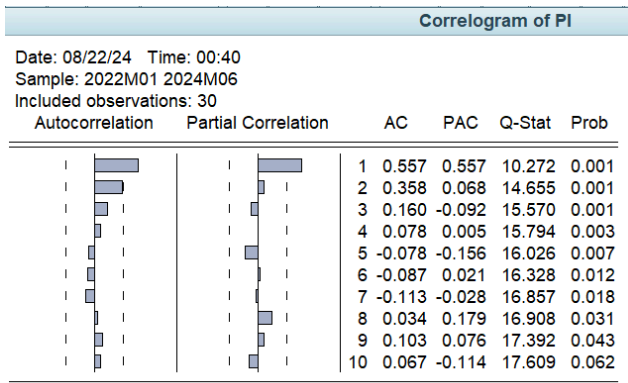


Figure 13. Results of the Correlogram test

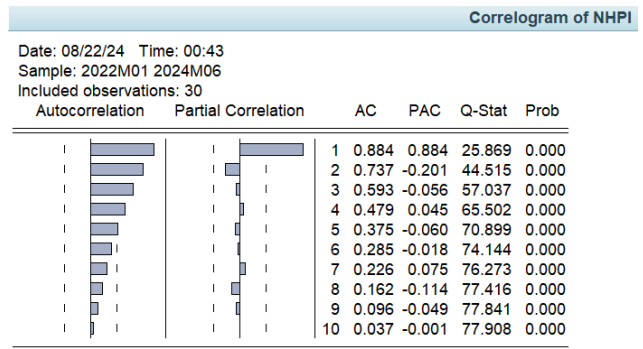


Figure 14. Results of the Correlogram test

2) Estimated Equation

Thus the model can be formulated as:

$$PI_t = \alpha + \beta_1 PI_{t-1} + \beta_2 NHPI_{t-1} + \beta_3 NHPI_{t-2} + \epsilon_t$$

where PI_t denotes the current sentiment index, PI_{t-1} denotes the prior period sentiment index, $NHPI_{t-1}$ denotes the prior period national housing prosperity index, $NHPI_{t-2}$ denotes the pre-previous national housing prosperity index, α is the intercept term, β_1 , β_2 and β_3 denote the coefficients of influence of the prior period sentiment index, the prior period national housing sentiment index and the pre-previous national housing sentiment index on the current period mass sentiment index, respectively, and ϵ_t is the random error term.

Using econometric software, we fitted the following equation as shown in Table 2:

Table 2. Estimated equation

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.477422	1.080937	0.441674	0.6640
PI(-1)	0.611872	0.166718	3.670114	0.0018
NHPI(-1)	7.588188	5.374828	1.411801	0.1751
NHPI(-2)	-7.902523	5.099202	-1.549757	0.1386
R-squared	0.473220		Mean dependent var	0.438768
Adjusted R-squared	0.385424		D. dependent var	0.063999
S.E. of regression	0.050172		Akaike info criterion	-2.983750
Sum squared resid	0.045310		Schwarz criterion	-2.785379
Log likelihood	36.82126		Hannan-Quinn criter	-2.937020
F-statistic	5.389963		Dubin-Watson stat	2.111339
Prob(F-statistic)	0.007975			

After statistical tests, we finalized the normal return equation as follows:

$$PI_t = 0.4774 + 0.6119PI_{t-1} + 7.5882NHPI_{t-1} - 7.9025NHPI_{t-2}$$

The specific meaning of the equation is that when the prior period sentiment index increases by 1 unit, the current sentiment index will increase by 0.6119 units; when the prior period national housing prosperity index increases by 1 unit, the current sentiment index will increase by 7.5882 units; when the pre-prior period national housing prosperity index increases by 1 unit, the current sentiment index will decrease by 7.9025 units. That is, the prior period sentiment index and the prior period national housing prosperity index play a positive role in current period sentiment, and the prior period national housing prosperity index plays a negative role in current period sentiment.

3) Possible Interpretations of the Equation

i. Positive effect of prior period sentiment index:

When the prior period sentiment index is high, this usually means that market participants are optimistic about the future. This optimism may continue into the next period because market participants' expectations and behavior tend to be inertial. Therefore, if the sentiment index is high in the prior period, the current period's sentiment may also be influenced by this positive sentiment and remain high.

ii. Positive effect of the prior period national housing prosperity index:

The National Housing Prosperity Index reflects the overall health of the real estate market, including indicators such as sales volume, price trends, and the number of new housing starts. If the national housing prosperity index is better in the prior period, this could mean that the real estate market is performing strongly, which may further boost the confidence of market participants. Positive real estate market performance may be seen as a sign of a healthy economy, which could boost investor and consumer sentiment.

iii. Negative effect of the pre-previous national housing prosperity index:

A poor national housing prosperity index in the pre-previous period may imply that the market has previously experienced a period of malaise. Even though the national housing prosperity index has improved in the prior period, the poorer performance of the prior period may still have a negative effect on sentiment in the current period. This may be due to the fact that market participants have a longer memory effect for negative information from the past, and even if the recent data has turned around, poorer conditions in the past may still affect their sentiment and decision-making.

Taken together, these findings suggest that market participants not only take recent information into account when forming their sentiment, but are also influenced by data from earlier periods. In addition, these findings may also suggest that, despite recent improvements in data, prior negative experiences may still dampen market optimism.

3.2.1.2. Calculation of abnormal returns and cumulative abnormal returns

1) Calculation of Abnormal Returns

Calculating abnormal returns means obtaining abnormal returns by subtracting normal returns from actual returns, which captures the likely impact of policy on data after the window period. The calculated equations are as follows:

$$AR = R_t - NR_t$$

Where R_t is the real return of the sentiment index, NR_t is the normalized return computed from the benchmark model. Substituting the 2024 data into the normal return equation yields NR_t ,

and finally AR is calculated by the abnormal return calculation equation. Specific values are given in Table 3:

Table 3. Abnormal return variance

	R	NR	AR
2024.1	0.5466	0.4309	0.1157
2024.3	0.5256	0.4846	0.041
2024.4	0.4710	0.5049	-0.0339
2024.5	0.5423	0.4724	0.0699
2024.6	0.5031	0.5192	-0.0161

2) Cumulative Abnormal Returns

Accumulate the abnormal returns to get the cumulative abnormal returns. The formula is as follows:

$$CAR = \sum_{i=t1}^{t2} AR_i$$

where t1 and t2 refer to the month in which the forecast period begins and ends, respectively. The final result of CAR is calculated to be 0.2509.

3.2.2. Results of the model and analysis of its possible reasons

When the results of an event study show a cumulative abnormal return greater than zero, this means that there was a positive abnormal return on actual returns relative to normal expected returns during the event window. This usually indicates a positive market reaction to the event.

In the real estate industry context, it means that if the market expects that the policy will be favorable to the real estate market, then consumers will have positive expectations about the extent of the real estate market boom after the policy is announced. At the same time, due to the lag in information dissemination, the public opinion sentiment did not drop sharply in the short term after the national housing prosperity index continued to decrease, but rather rose and then fell, which also proves the positive impact on the public sentiment brought by the favorable policy from another perspective.

Another point that corroborates the lagged effect of the above news is that, looking at the abnormal return spreads for each month, we see that the spreads are gradually shrinking, i.e., the sentiment of public opinion is gradually reverting to a normal return pattern over time, and the good expectations brought about by the favorable policies are declining, or even reversing into a more negative sentiment.

Although we find that the sentiment of public opinion may ultimately be determined by the actual situation of the domestic real estate market (e.g., by referring to the national housing prosperity index), favorable real estate policies do have a relatively significant effect on improving public sentiment. This suggests that public confidence can be boosted through the release of positive policies, which may have a positive effect on improving the status quo of the industry.

4. Conclusion

4.1. Research review

The purpose of this study is to explore the impact of real estate policy changes on the sentiment of popular opinion. Firstly, we crawled the pop-ups and comments of Bilibili and Xiaohongshu through python crawler technology, and then we calculated the sentiment index by citing the SnowNLP library for sentiment analysis. After collecting the national housing prosperity index and mass opinion sentiment index from January 2022 to June 2024, we analyzed the decision equation of the mass sentiment index before the introduction of the policy by constructing a dynamic panel model and paid special attention to the role of the pre-previous mass sentiment index and the pre-previous national housing prosperity index on the current sentiment.

Next, we built an event study model to analyze whether the introduction of the policy in the first quarter of 2024 had an impact on the sentiment of public opinion in the second quarter by calculating the difference between the expected sentiment index and the actual sentiment index and figuring out the cumulative abnormal return difference.

4.2. Main findings

The results of the equation fitting indicate that the pre-previous sentiment index and the pre-previous national housing prosperity index play a positive role in the current period's sentiment, while the pre-previous national housing prosperity index negatively affects the current period's sentiment. In addition, the results of the event study show that the cumulative abnormal returns are significantly greater than 0 after a change in real estate policy, which suggests that the market reacts positively to a change in real estate policy. That is, the popular sentiment index will improve after favorable and positive policies.

4.3. Significance of the research

This study provides empirical evidence for understanding how real estate policy changes affect the sentiment of public opinion. The findings have important implications not only for policymakers in the real estate industry but also provide valuable insights for investors and related stakeholders. In addition, the findings suggest that changes in sentiment indices are closely related to the health of the real estate market, which has important implications for monitoring market sentiment and formulating related policies.

4.4. Limitations and future research directions

While this study provides some valuable insights, there are some limitations. First, the study only considered the impact of a few key variables, and future research could consider more factors, such as economic policy uncertainty and political stability. Second, the study does not point out in depth the specific mechanism of policy influence on sentiment indices. In addition, future research could explore the differences in sentiment responses in different regions or different types of real estate markets.

4.5. Concluding remarks

In summary, this study provides strong evidence for understanding the impact of real estate policy changes on the sentiment of public opinion. These findings help policymakers better assess the effects of policies and provide investors with insights about changes in market sentiment. Future research will further deepen our understanding of the complex relationship between real estate markets and sentiment responses.

References

- [1] LIU, Ji. (2023). The Influence of Real Estate Market Sentiment on China's Housing Price Volatility: Based on the Perspective of Economic Policy Uncertainty. doi: <https://doi.org/10.27162/d.cnki.gjlin.2023.002495>.
- [2] ZHANG, W., LI, Y. and WANG, P. (2022). Social Media and Capital Markets: A Bibliometric Analysis. *China Journal of Econometrics*, 2(01), pp.32–57.
- [3] HUANG, Y., HONG, W. and YU, H. (2019). HOW DOES THE MARKET SENTIMENT AFFECT URBAN HOUSING PRICE. *Economic Theory and Business Management*, (07), pp.75–88.
- [4] Zhuang Jinying (2003). Relationship of Emotion and Decision Making. *Advances in Psychological Science*, 11(04), pp.423–431.
- [5] Zhao, W., Geng, Y. and He, Y. (2018). Media Coverage, Expectation and Housing Price Volatility. *Economic Review*, (02), pp.133–146. doi: <https://doi.org/10.19361/j.er.2018.02.10>.
- [6] KUANG, W., ZHANG, S. and ZHANG, Z. (2020). MEDIA SENTIMENT AND HOUSING PRICE VOLATILITY. *Economics Theory and Business Management*, (12), pp.56–58.
- [7] JIANG, F., MENG, L. and TANG, G. (2021). Media Textual Sentiment and Chinese Stock Return Predictability. *China Economic Quarterly*, 21(04), pp.1323–1344. doi: <https://doi.org/10.13821/j.cnki.ceq.2021.04.10>.
- [8] WANG, J. (2006). The Study on the Evolution of Business Cycle Theories and Schools. *Journal of Capital University of Economics and Business*, (03), pp.23–28.