HONG KONG MONETARY AUTHORITY

Research Memorandum 05/2025

28 January 2025

A GENERATIVE ARTIFICIAL INTELLIGENCE APPROACH TO TRACKING MAINLAND CHINA'S HOUSING MARKET SENTIMENT USING SOCIAL MEDIA DATA

Key points

- *To enhance monitoring of Mainland China's housing market, this paper develops a daily index to track the public sentiment in the sector using Chinese social media data and generative Artificial Intelligence (GenAI). After web-scraping granular textual and video data from a major Chinese social media platform (Sina Weibo), we utilise an in-house GenAI powered by a large language model that specialises in Chinese tasks to pre-process the data and assess housing market sentiment.*

- *Throughout this process, we keep human-in-the-loop by comparing samples of machine results with human evaluations and refining our prompts. We find that the assessments made by the GenAI are comparable to human evaluations. To reduce potential bias in our sentiment index, we also develop strategies to identify social bots and cyber trolls and eliminate microblogs posted by them.*

- *Our analysis reveals that the GenAI-driven sentiment index effectively captures trends in public sentiment on the housing market, particularly around landmark events, and possesses predictive powers for property sales. In addition, we go beyond the national housing market sentiment index, leverage the strong comprehension capability of GenAI to identify the cities being discussed in microblogs, and develop more granular sentiment indices at the city-tier level or provincial level. Again, our city level sentiment index is helpful to predict local property sales. These findings represent a novel contribution to the literature.*

- *Given the importance of the real estate sector to Mainland China's economy, our research should help policymakers monitor the property market in a timely manner and at a new level of granularity. Our empirical results also highlight the transformative potential of GenAIs in research and macroeconomic surveillance, as we can now tackle tasks that were previously deemed infeasible, establishing GenAI as a valuable and promising option for applications in big and unstructured data.*

Prepared by: Zhang Wu, Michael Cheng, Philip Ng, and Alice Wang[1]
Economic Research Division, Research Department
Hong Kong Monetary Authority

[1] The authors would like to thank Andy Cheung and Iris Yin for their excellent research assistance, and express our gratitude to Max Kwong, Oscar So, and Maggie Poon for their valuable technical support. SARA, the Hong Kong Monetary Authority's in-house generative Artificial Intelligence application, helps proofread this paper.

I. INTRODUCTION

1. Thanks to significant advancements in neural networks and computational power in recent years, generative Artificial Intelligence ("GenAI" hereafter), which generates responses based on input data and appropriate prompts, has achieved cutting-edge performance in various natural language processing tasks, including information summarising and question answering. Unlike traditional sentiment analysis methods that rely on pre-defined dictionaries (Tetlock (2007) and Loughran and McDonald (2011)), recent research suggests that GenAI can accurately assess a word's sentiment value within a specific context[2], although the sentiment conveyed by a word may vary significantly in different contexts (Ruan et al. (2020)).[3] Besides accuracy, assessments by GenAI are more efficient than human evaluations and traditional machine learning methods used in the literature that are based on manual labelling (Liu (2010)). In fact, these older practices are time-consuming, error-prone, and impractical on a large scale; conversely, GenAI offers a new way that can help address previously intractable problems (BIS (2024a)).

2. Utilising an in-house GenAI (**S**ecured **AI** **R**esearch **A**ssistant, "**SARA**") that specialises in Chinese tasks, this paper develops a daily sentiment index that tracks public sentiment in Mainland China's housing market based on social media texts and videos, with the aim of enhancing the monitoring of the sector in a comprehensive and timely manner.[4] Specifically, we first construct a GenAI-driven housing market sentiment index starting from 2013, following several necessary steps including the elimination of identified social bots and cyber trolls. Next, we demonstrate that the index serves as a leading indicator for property sales and outperforms the conventional lexicon-based approach. Furthermore, we extend our analysis beyond the national sentiment index and develop housing market sentiment indices at a more granular level, leveraging the strong comprehension and reasoning capability of the GenAI to identify localities.

[2] GenAI has been shown to outperform conventional models in understanding the tones of central bank communications (Hansen and Kazinnik (2023)) and in predicting stock returns based on sentiment analysis of news headlines (Lopez-Lira and Tang (2023)).

[3] For example, the sentiment conveyed by the following sentences cannot be assessed merely by counting the number of positive and negative words in each: "I enjoy it", "I wish I could enjoy it", and "I cannot say that I do not enjoy it".

[4] Indeed, one important application of GenAI for central banks today is nowcasting, which helps monitor developments in the real economy in a timely manner (BIS (2024b)).

In fact, such granular and informative housing market sentiment indices in Mainland China are absent in previous studies, and our findings represent a novel contribution to the literature.[5] From a broader perspective, our research confirms the transformative potential of GenAI in macroeconomic surveillance and research, particularly with alternative data, and demonstrates GenAI's capability to provide insights into previously impractical tasks, thereby enhancing the understanding of important issues.

3. The rest of the paper is organised as follows. Section II introduces the data and methodology, with emphasis on our experience in prompt engineering and validation, as well as the identification of social bots and cyber trolls. Section III presents the empirical results. In particular, we rely on the GenAI to construct the national housing market sentiment index and examine its correlation with real economic activities. We also utilise the GenAI to identify the cities being discussed in microblogs, aiming to develop a local-level sentiment index for the property market. Section IV looks into the related policy implications and provides some concluding remarks.

II. DATA AND METHODOLOGY

2.1 Data and sample

4. To construct the property sentiment index, we webscrape microblogs from Sina Weibo ("Weibo" hereafter), a major social media platform in Mainland China.[6] Unlike Shao et al. (2023) and Zhu et al. (2023), which limit microblog samples to those that include the keyword of "housing price" only, we select posts based on a range of terms that match the relevant topics (see Annex A). Specifically, based on the list used by Li

[5] In the literature that develops a city-level sentiment index, one strand extracts sentiment from market-based indicators (Ding et al. (2023) and Zhou (2018)). These studies employ market proxies to indirectly generate a sentiment index, which may exhibit a time lag. Another strand uses geographic information of the texts such as the registration city of the user and IP address (Zhu et al. (2023)). In practice, however, this information may often be absent. As far as we are aware, this is the first analysis that constructs housing market sentiment indices across all tiers of cities for Mainland China using GenAI. This method can yield a greater number of outcomes in a more timely and accurate manner.

[6] Weibo was ranked third among Chinese social media platforms in 2021 by a marketing agent based on publicly available data. According to Weibo's financial report as of the third quarter of 2024, its monthly active users reached 587 million in September 2024, up from 573 million in December 2021. Also, it has an average daily active user count of 257 million in September 2024.

et al. (2022) and our own experience, we identify microblogs related to Mainland China's property market using (i) certain hashtags that the authors add to indicate the intended themes, and (ii) super topics[7]. In addition, we remove duplicated records through a unique microblog ID. We download the identified microblogs from January 2013 to December 2024, encompassing as many cycles in the Mainland real estate sector as possible.

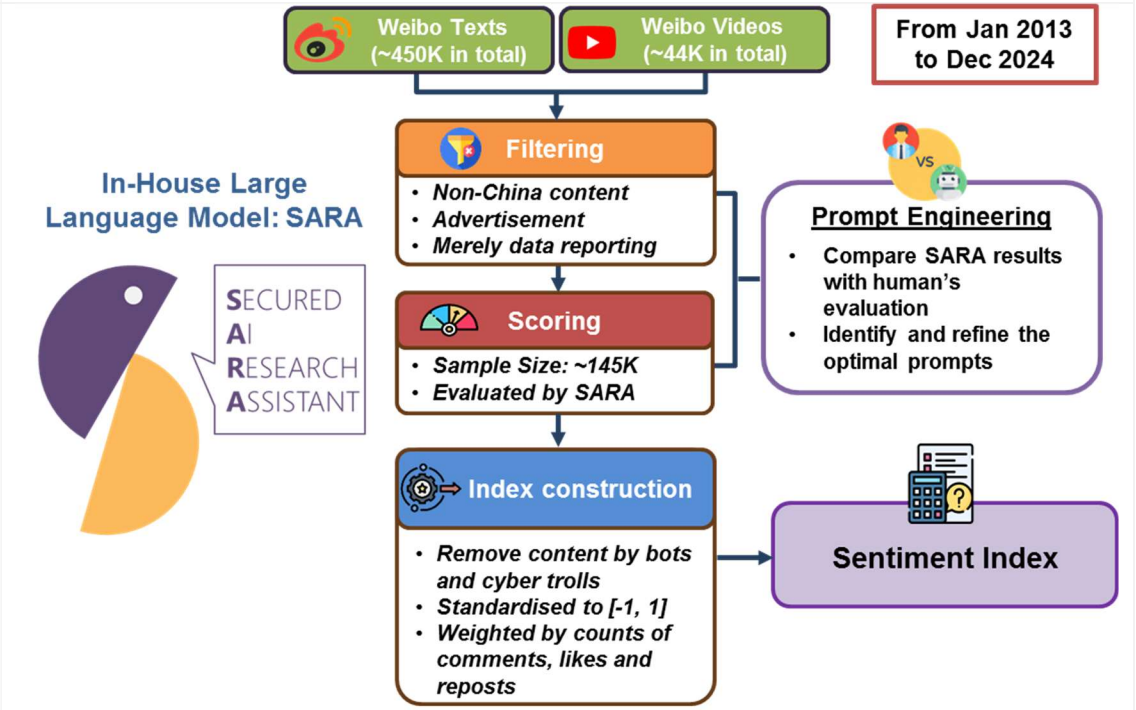
5. A prominent trend in social media nowadays is the rise of short videos. These videos have gained immense popularity due to their engaging nature and their capability to convey information quickly. With advancements in multimodal models that are capable of seamlessly handling various input formats, we not only include traditional textual elements of microblogs but also incorporate visual media from posts. To be specific, we convert these videos to texts using the speech-to-text function of our in-house GenAI. In total, our raw dataset contains more than 450,000 textual microblogs themed on the property market, along with nearly 44,000 videos.

2.2 The construction of the property sentiment index

6. The construction of the GenAI-driven property sentiment index consists of three key steps: (i) filtering, (ii) scoring, and (iii) index construction. The first two steps rely on our in-house GenAI, which is powered by a large language model that is specially designed for Chinese tasks as well as our needs, and hosted entirely on-premises to ensure security, privacy, and control. Considering that the original microblogs are in Chinese, we use Chinese as the input and instruction language to ensure that the GenAI understands the target content accurately without any information loss. The entire procedure is visualised in Chart 1.

[7] Super topics are a distinctive feature on Weibo. Both topics and super topics serve to categorise content and facilitate discussions among users. However, unlike regular topics, super topics have dedicated pages for community engagement and structured interactions, resulting in a more organised experience for users.

Chart 1: The development process of the GenAI-driven housing market sentiment index



Note: SARA is the name of our in-house GenAI.

Source: HKMA staff illustration.

7. Filtering. We first filter out microblogs that are either irrelevant to Mainland China's property market or unable to reflect sentiment. For example, some microblogs from official accounts[8] or news agencies are merely data reporting without conveying public sentiment, while others discuss cases from other economies or the non-property sectors in Mainland China (see examples in Annex B). Leveraging our in-house GenAI model, which is able to identify the core message of texts and capture subtle nuances, we design prompts instructing the model to remove posts that (i) contain non-China or non-property market content, (ii) belong to advertisements, and (iii) merely report or repeat official data. On top of that, we also delete posts from official accounts and news media that we have identified manually. This initial filtering process is crucial to maintain a high-quality dataset that accurately reflects the public sentiment in Mainland China's property market.

[8] Official accounts are run by either the state or the owning company of Weibo (i.e. Sina). Examples of these accounts include Sina Finance, China Daily, People's Daily, and Xinhua. Such accounts likely post microblogs that merely report data or the official stance. As our focus is on the public opinion, official accounts are thus ruled out.

8. Scoring. After excluding the unrelated posts, we have a dataset of around 145,000 filtered microblogs, which forms the foundation for our further analysis. Using the in-house GenAI, we ask the model to evaluate the sentiment score for each microblog in the dataset on a scale of 0 to 10, with a score of 10 being the most optimistic and 0 being the most pessimistic. This granular rating system enables us to capture a wide range of sentiments expressed by users.

9. Index construction. After scoring, we remove microblogs posted by suspicious social bots and cyber trolls to reduce sentiment bias (more details are provided in the discussion below).[9] We construct the sentiment index taking into account the potential influence of each microblog. In particular, sentiment scores at the microblog level are standardised to a continuous range of -1 to 1, where -1 indicates extreme pessimism, 0 represents a neutral tone, and 1 refers to extreme optimism. Given the varying influence of each microblog, we then calculate the daily average of sentiment scores, weighted by the number of influential factors (i.e. likes, comments, and reposts) as shown in the following equation.[10]

$$Daily\ \ sentiment\ \ index_t = \sum_{i=1}^{n} [\frac{(w_1 * Comment\ Count_i + w_2 * Like\ Coun_{\ i} + w_3 * Repost\ Coun_{\ i} + 1)}{\sum_{i=1}^{n} w_1 * Comment\ Coun_{\ i} + w_2 * Like\ Count_i + w_3 * Repost\ Count_i + 1} \times Sentiment\ Scor_{\ i}]$$

2.3 Prompt engineering and validation

10. Considering that prompt engineering plays a crucial role in optimising the performance of GenAI and enhancing the quality of output (Chen and Zhao (2024)), we apply this optimisation approach to our "filtering" and "scoring" steps, as well as to other instructions or prompts where applicable. Based on the literature and our own experiments, we follow these four principles to design and refine our prompts: (i) commands must be detailed and specific, (ii) the GenAI is instructed to take on a specific role,

[9] To reduce the direct impact of social bots and cyber trolls that may lead to extreme tones, we identify the relevant users based on several rules and exclude their posts accordingly to ensure a spam-free sentiment index.
[10] The weights of the influential factors, w_1, w_2, and w_3, are set at the values of five, three, and one, respectively, to reflect the relative importance of comments, likes and reposts. Our major findings are robust to slight variations in the values of these weights. In case all influential factors are zero, we replace the zeros with ones in both the numerator and the denominator.

(iii) an iterative approach is employed to refine prompts based on the GenAI's responses, and (iv) a chain-of-thought approach is adopted to break down a problem into smaller, manageable parts and the GenAI then reasons through them step by step.

11. In the process, we also emphasise a human-in-the-loop approach to ensure the accuracy and reliability of our findings, by comparing the "filtering" and "scoring" results generated by the GenAI with human evaluations on randomly selected samples (600 posts for each step). Additionally, to ensure the accuracy and consistency of human results, each microblog is assessed by two reviewers.

- First, the filtering results from humans and the GenAI are compared, and the comparison results are presented as a confusion matrix in Chart 2A. The matrix indicates good classification performance for the GenAI, with a high recognition rate for microblogs that it successfully identifies for both removal and retention.[11]

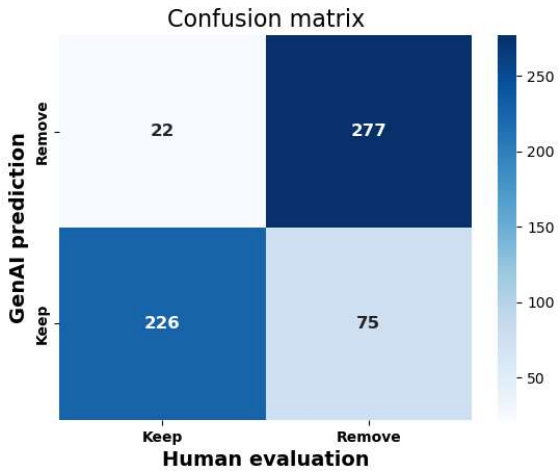
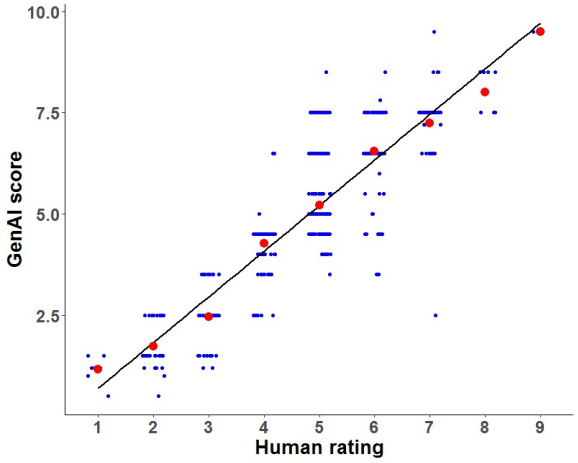
- Second, we compare the sentiment scores generated by the GenAI and humans for each microblog and present the results in Chart 2B, using a method similar to Shapiro et al. (2022). The chart shows that the scores evaluated by the GenAI are, on average, strongly and positively correlated with those of humans, demonstrating that the scores from the GenAI and humans are very similar.[12]

Overall, these results suggest that the prompts we provided enable our in-house GenAI to accurately identify microblogs relevant to our research focus and score the texts precisely.[13]

[11] In fact, the overall accuracy (i.e. the proportion of correctly identified results among the total number of cases examined) is about 84% in our sample. In terms of the Macro-F1 statistic, a widely recognised standard in the literature for assessing model classification performance in the field of sentiment analysis, our in-house GenAI also achieved a high score of 85.1% in our sample.

[12] The Spearman rank correlation for the scores by the GenAI and humans is 0.83.

[13] To illustrate that our in-house GenAI performs the scoring task well, some examples are listed in Annex C.

Chart 2A: GenAI and human filtering results comparison	Chart 2B: GenAI and human rating results comparison
	
Note: The labels of "Remove" on the x- and y-axes indicate that a microblog is deemed to be ruled out by either GenAI or humans, while the labels of "Keep" indicate the opposite.	Note: Blue dots are raw data points, solid line represents linear regression fit line, and red points are within-rating means of raw data.
Sources: Weibo, in-house GenAI and HKMA staff calculations.	Sources: Weibo, in-house GenAI and HKMA staff calculations.

2.4 Social bots and cyber trolls

12. Previous studies have documented that public opinions can be significantly influenced by misinformation and extreme sentiments propagated through social media. To mitigate the risk of potential bias on our sentiment index stemming from social bots and cyber trolls, we employ three methods to identify these two types of microblog publishers based on the filtered data, as suggested by Gorodnichenko et al. (2021). First, we calculate the total number of microblogs published on the day a new account is established and flag an account as a potential cyber troll if the number exceeds five, considering that cyber trolls tend to create new accounts and use them to convey specific messages. Second, we add up the number of microblogs published during abnormal periods for each user, and flag an account as a potential bot if it posts at least five microblogs between 2 AM and 6 AM on any given day, as this timeframe is typically inactive for human users. Third, we compute the number of identical microblogs at the user level, and flag an account as a bot and/or cyber troll if the repeated messages are published within a relatively

short time interval (i.e. five seconds).[14]

13. Following these rules, we find that there is limited evidence to support a notable effect of social bots and cyber trolls on our sentiment index, with only 11 out of approximately 25,000 users identified as target entities. This finding partly demonstrates the effectiveness of our previous filtering process, which excludes unrelated posts, such as advertisements that resemble those posted by bots and internet trolls. As a result, we have eliminated all microblogs (about 8,200) posted by the identified users before constructing our final sentiment index.

III. EMPIRICAL RESULTS

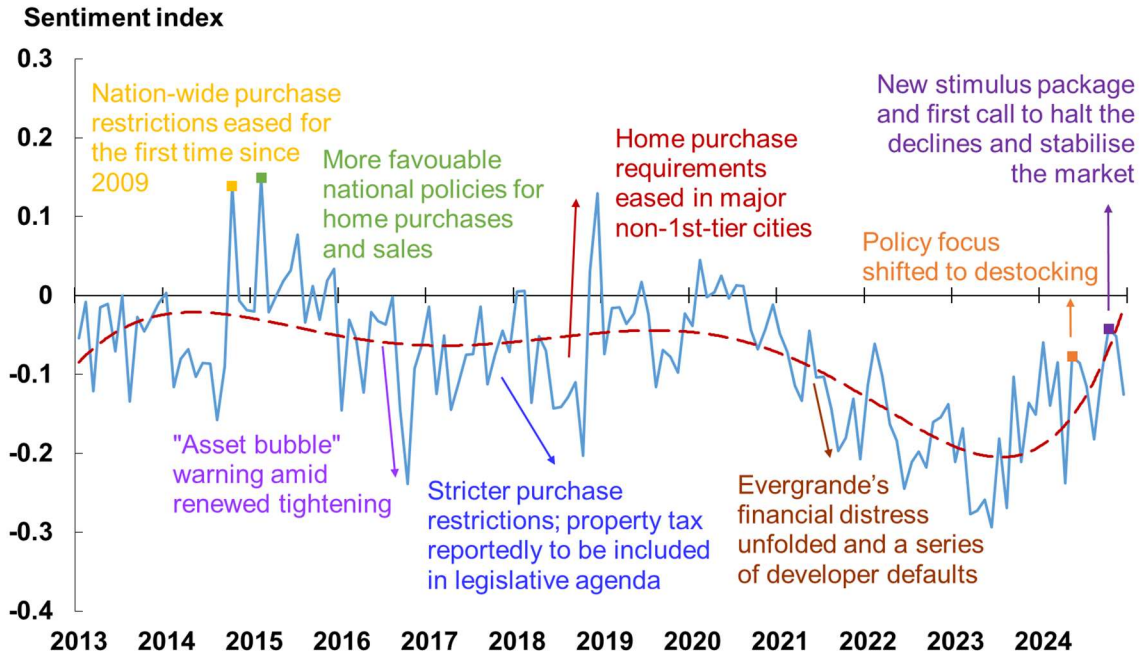
3.1 Measuring property market sentiment in Mainland China

14. Chart 3 displays the monthly GenAI-driven sentiment index for Mainland China's housing market, with several historical landmark events highlighted. The chart illustrates that the index has effectively captured shifts in sentiment in response to landmark events and exhibited cyclical patterns over the past decade, partly reflecting the dynamics of the housing cycles. Specifically, the index improves notably amid favourable events such as major policy easing. For example, in late 2014 and early 2015, the sentiment index surged following the first nationwide easing of purchase restrictions since 2009 and favourable policies to boost housing transactions.[15] On the other hand, the sentiment index declined in 2021 in response to potential drivers such as the rising default risks from certain developers. In the latest episode, our index suggests some improvement in housing sentiment following the rollout of the new stimulus package by the authorities since late September 2024, but saw some consolidation near the end of 2024.

[14] Since our sample does not include reposted microblogs, copying and pasting the text after viewing a published microblog takes significantly more time than merely reposting. Therefore, the five-second time lag between the publication of the first microblog and another user's reaction is considered a reasonable threshold for human users, distinguishing them from cyber trolls and bots.

[15] Such relaxations of purchase restrictions include lowering down payment ratios, easing tax rules for home sales, and trimming mortgage rates.

Chart 3: GenAI-driven housing market sentiment index for Mainland China

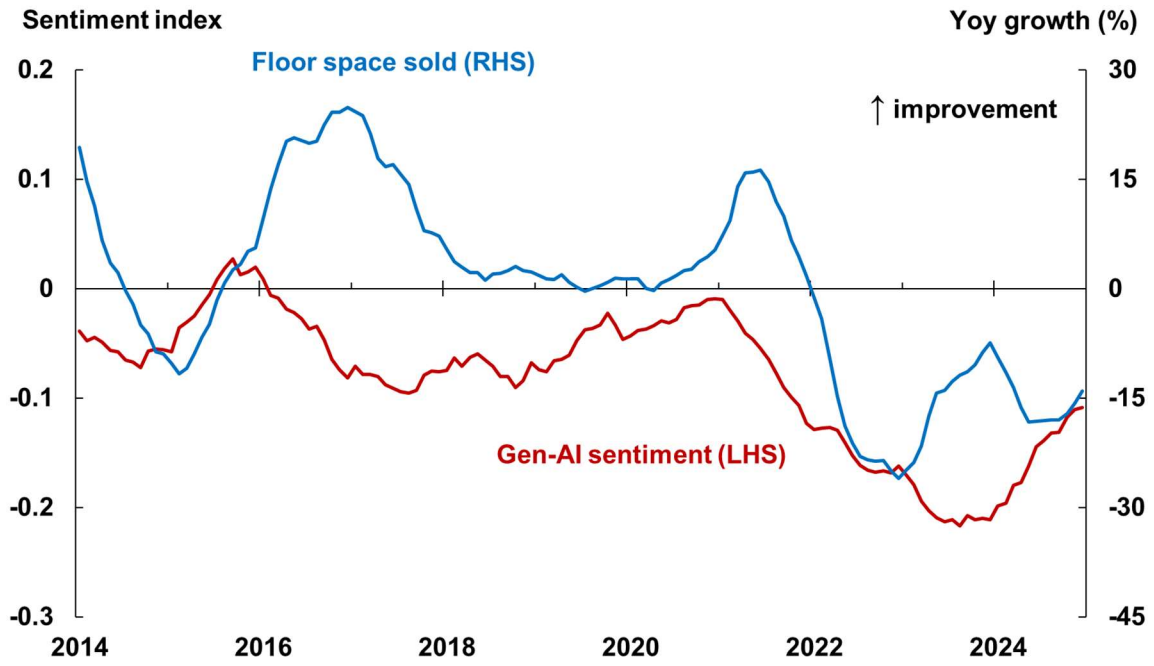


Notes: Data are converted into a monthly frequency. The red dashed line indicates the 12-month moving average of the sentiment index.
Sources: Weibo, in-house GenAI and HKMA staff calculations.

15. For surveillance purposes, we find that our GenAI-driven index is a leading indicator of property market activities. Chart 4 compares our sentiment index with the growth of floor space sold[16], and indicates that the smoothed sentiment index appears to have led the floor space sold. Indeed, the Granger causality test confirms that our GenAI-driven sentiment index leads floor space sold by one month, with no reverse relationship identified. This demonstrates that our sentiment index possesses predictive power for floor space sold, thereby facilitating the monitoring of property market activities and the identification of their turning points.

[16] As a measure of the property sales, floor space sold indicates housing demand and affects developers' liquidity.

Chart 4: GenAI-driven housing market sentiment index and national floor space sold



Note: The two indicators are smoothed using a 12-month moving average to enhance comparability.

Sources: CEIC, Weibo, in-house GenAI and HKMA staff calculations.

3.2 Comparison with news-based sentiment index

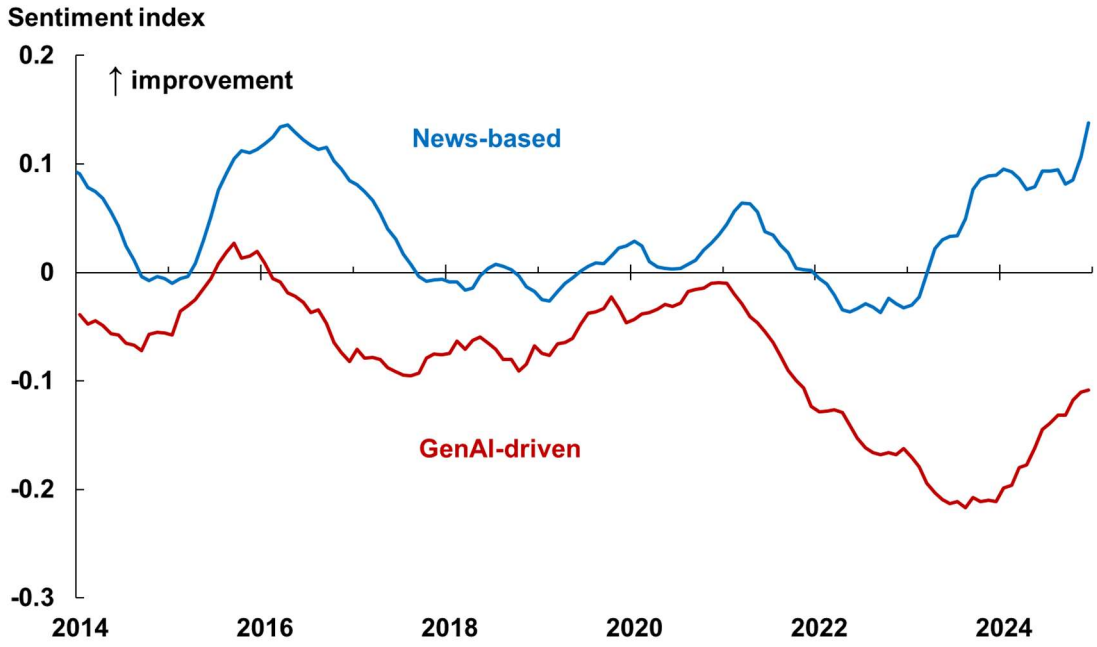
16. Compared with those sentiment indices generated by traditional textual analysis methods, the GenAI can be a superior evaluation tool for gauging the public sentiment on the property market. We plot the sentiment indices estimated by both methods in Chart 5, where the news-based property market sentiment index is generated by a traditional lexicon-based approach through counting positive and negative words in each news article.[17] One interesting observation is that the two indices show co-movement. However, the news-based one generally conveys a more positive tone, while the GenAI index conveys a more cautious tone in general. We attribute this difference in part to the underlying data sources.

17. Further analysis indicates that the GenAI-driven index serves as a more effective leading indicator for property market activities compared to its news-based

[17] For more details, see "Box 2: A news-based property market sentiment index for Mainland China" in the Hong Kong Monetary Authority's Half-Yearly Monetary and Financial Stability Report (September 2021 issue).

counterpart. The Granger causality test suggests that the news-based sentiment index adds only marginal value in predicting floor space sold. This result, along with the superior representation of public attitudes, suggests that the GenAI-driven index can be a better proxy for the public sentiment on the property market.

Chart 5: GenAI-driven vs. news-based property market sentiment indices



Notes: The news-based property market sentiment index is computed as the standardised difference between the share of positive words and that of negative ones in news articles; while the GenAI-driven property market sentiment index is generated by the in-house GenAI after evaluating the content of social media. Both series are smoothed using a 12-month moving average to enhance comparability.

Sources: Weibo, Wisers, in-house GenAI and HKMA staff calculations.

3.3 Regional GenAI-driven sentiment index

18. Apart from a national sentiment index, we have also developed regional sentiment indices, leveraging the strong comprehension and reasoning capability of the GenAI. We believe that these granular sentiment indices represent a novel contribution to the literature. While some previous studies (e.g. Zhu et al. (2023)) employed geographic information such as the registration city of user account and IP address as a proxy for cities, our method is based on the content of the post, and we believe that our approach has several advantages. Firstly, the geographic information (i.e. city name

derived from IP address) is scanty in our sample[18]; secondly, even if such information is provided, the city shown is not necessarily the major city mentioned or talked about in the microblog.[19] Therefore, employing content-based methods to identify the primary location in the post is more accurate and pertinent, and is able to yield more data.

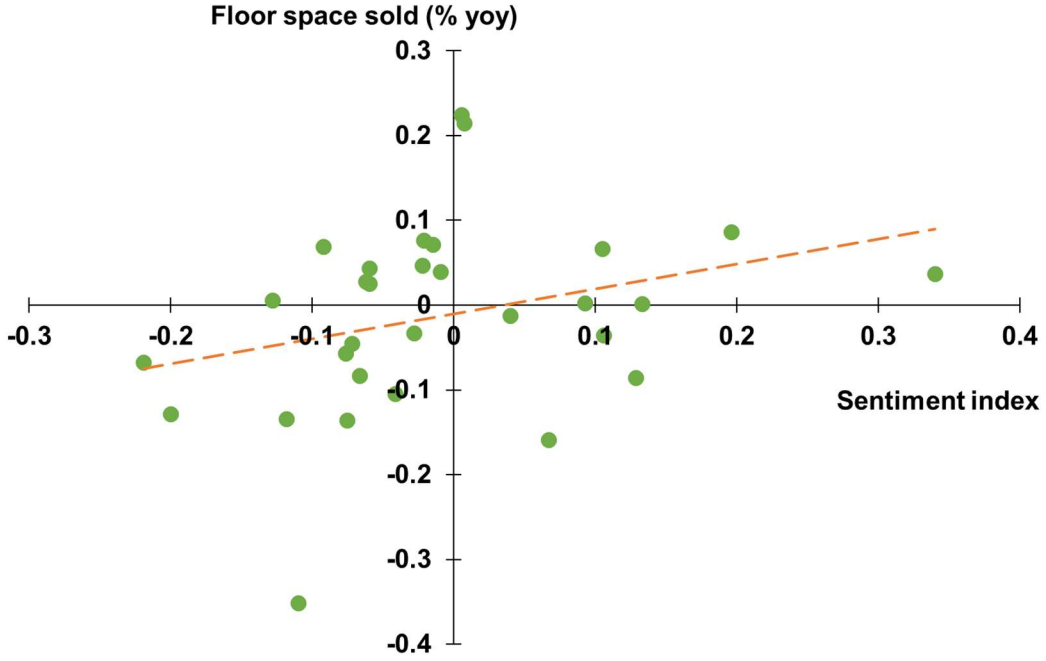
19. In fact, among the content-based approaches to identifying cities, the GenAI approach is significantly more powerful than traditional techniques that extract city names from a predetermined list of selected cities. We find that the GenAI not only can recognise the city in focus even when multiple cities are mentioned, but also can infer the appropriate cities even if the city names are not explicitly stated. For example, social media users often mention one of a city's districts or a famous scenic spot instead, or to use the city's nickname (e.g. "Imperial City" for Beijing, "Magic City" for Shanghai).[20] Under these circumstances, the state-of-the-art GenAI can effectively handle such cases well, despite the presence of noise in the posts.

20. Taking advantage of the powerful capabilities of the GenAI, we identify the cities being discussed in microblogs and compile a city-level index. In particular, we instruct the GenAI to identify the prefecture-level city discussed or implied[21], if any, in each microblog and compile the sentiment indices at the city-month level. This process generates a dataset comprising about 60,000 microblogs with identified city information.

21. Considering the evidence that the national GenAI-driven sentiment index effectively predicts national floor space sold one month in advance, we further investigate whether this relationship holds at the city level. Specifically, we plot the growth of each city's floor space sold alongside its average sentiment during our sample period, with sentiment scores calculated with a one-month lag for each city. Chart 6 demonstrates that the leading sentiment is positively correlated with regional property sales.

[18] IP addresses are required to be provided by the Mainland authority since May 2022. However, many of them are still missing in our sample even after 2022.
[19] Affuso and Lahtinen (2019) also mentioned that measures relying on IP addresses are less precise.
[20] To illustrate the capability and performance of the in-house GenAI, some examples of the city names identified are provided in Annex D.
[21] If multiple cities are identified as eligible in the microblog, the sentiment expressed in the post will be assigned to each city.

Chart 6: Correlation between property sales and sentiment index at the city level



Notes: Each dot represents the average floor space sold and the average sentiment in a city during the sample period, while the orange line denotes the fitted trend. Sentiment is calculated one month behind the floor space sold.

Sources: Wind, CEIC, Weibo, in-house GenAI and HKMA staff calculations.

22. To formally test the finding, we use a city-month level panel data to run the regression, controlling for city-level characteristics, as well as city and time fixed effects.

$$Sold_{i,m} = \alpha_i + \mu_{y,m} + \beta Sentiment_{i,m-1} + \gamma Price_{i,m-1} + \delta X_{i,y-1} + u_{i,m}$$

where $Sold_{i,m}$ is the year-on-year growth rate of floor space sold for month m in city i, $Sentiment_{i,m}$ refers to our GenAI-driven sentiment index for month m in city i, $Price_{i,m}$ refers to the growth of property price for month m in city i, α_i and $\mu_{y,m}$ control for city and year-month fixed effects, respectively. The growth of floor space sold and sentiment are winsorised at the 5% level. In addition, we include $X_{i,y}$ to control for city-specific characteristics for year y to which m belongs. The variables used are annual GDP growth and urbanisation ratio. β is the coefficient of interest and it measures the effects of our sentiment index on the property sales. The estimation results are shown in Table 1.

Table 1. Regression results of sentiment index on property sales at the city level

	Sold (% yoy)			
	(1)	(2)	(3)	(4)
Sentiment	0.036***	0.037***	0.037***	0.034***
	(0.013)	(0.013)	(0.013)	(0.014)
Price	N	N	Y	Y
Annual control variables	Y	Y	Y	Y
City effect	N	Y	N	Y
Year-month effect	Y	Y	Y	Y
No. of cities	59	59	41	41
Observations	1,756	1,756	1,592	1,592
R-sq.	0.011	0.049	0.011	0.046

Note: Robust standard errors clustered at the city-level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Sources: Wind, CEIC, in-house GenAI and HKMA staff calculations.

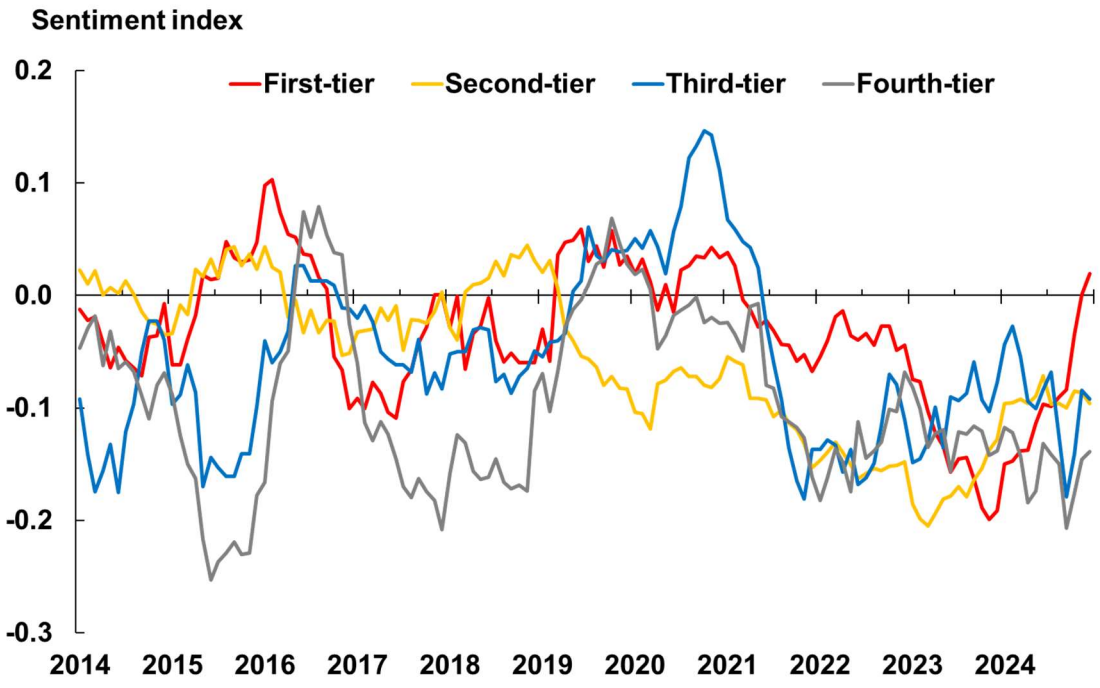
23. In general, the results suggest that our GenAI-driven sentiment index significantly enhances the prediction of the property sales in each city one month in advance. After accounting for both city and time fixed effects, as shown in Column (4), a one standard deviation increase in the sentiment index (equivalent to 0.318), on average, results in a 1.08-percentage-point rise (0.318 x 3.4%) in the growth rate of the floor space sold in that city the following month. The results are little changed in the robustness check through adjusting fixed effects.

24. To gain more insight at the city-tier level, we further construct four city-tier level indices, where the classification of each city is primarily based on the official system.[22] As shown in Chart 7, our sentiment indices display similar cyclical patterns across various city tiers. However, the indices also exhibit disparities during specific periods, partly reflecting different reactions to landmark events. For instance, the sentiment indices diverged in early 2015, with first- and second-tier cities experiencing an uptick in sentiment driven by nationwide easing policies, such as lower down payment ratios and more tax benefits. In contrast, lower-tier cities saw a decline in sentiment, largely due to their high inventory-to-sales ratios. After mid-2015, the sentiment in third- and fourth-tier cities rose sharply,

[22] According to the National Bureau of Statistics (NBS), the first-tier cities are Beijing, Shanghai, Guangzhou, and Shenzhen, the second-tier cities refer to 31 cities, while the third-tier cities indicate 35 cities (see the website of the NBS for the full list). Based on the list, we define all other cities in our sample as fourth-tier cities.

coinciding with the large-scale implementation of monetised resettlement for urban renewal and varying city-specific policies that reportedly favoured lower-tier cities the most. In late 2024, the sentiment in first-tier cities rebounded following a comprehensive package of policies, while the sentiment in lower-tier cities remained somewhat sluggish.

Chart 7: GenAI-driven housing market sentiment index by city tier



Note: All series are smoothed using a 12-month moving average to facilitate visualisation.
Sources: Weibo, in-house GenAI and HKMA staff calculations.

25. Additionally, we prompt the GenAI to group the cities into their corresponding provinces and calculate the average sentiment scores within the same province. This data structure helps us visualise the comparison of sentiment changes across regions after major policy introductions. According to our results, among the provinces for which data is available, overall local sentiment became more optimistic in October 2024 compared to April 2024, especially in Beijing and Shanghai. The improvement was caused by a new round of policy support measures, exemplified by the removal of purchase restrictions in the first-tier cities. As illustrated in Charts 8A–8D, one interesting observation is the worsening of sentiment in Hebei, contrasted with the enhancement of sentiment in Beijing, partly reflecting the siphoning effect of the capital city on surrounding cities. That said, the phenomenon does not appear in provinces surrounding Shanghai,

where the public believes economic developments are more even across regions.

Chart 8A: Sentiment in provinces around Beijing: Before the latest round of supportive measures (April 2024)

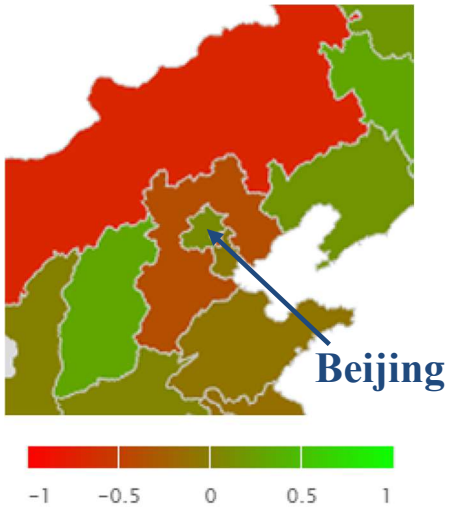


Chart 8B: Sentiment in provinces around Shanghai: Before the latest round of supportive measures (April 2024)

Chart 8C: Sentiment in provinces around Beijing: After the latest round of supportive measures (October 2024)

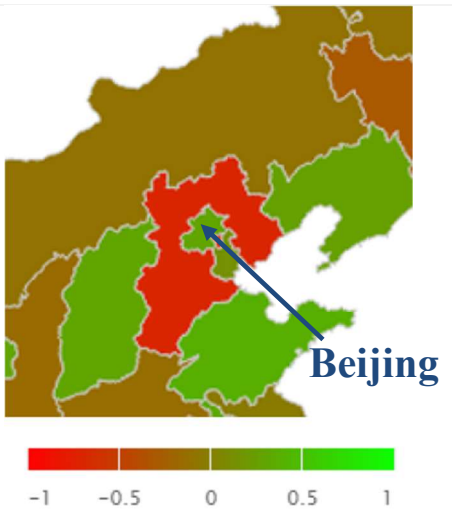


Chart 8D: Sentiment in provinces around Shanghai: After the latest round of supportive measures (October 2024)

Notes: Missing values are replaced by the provincial sentiment from the previous month.

Sources: Weibo, in-house GenAI and HKMA staff calculations.

IV. POLICY IMPLICATIONS AND CONCLUDING REMARKS

26. To conclude, our GenAI-driven sentiment indices based on social media data effectively capture the public sentiment in Mainland China's property market, successfully tracking the housing market's response to major policies and market events. By providing more granular sentiment indices at the city level or provincial level, we can also observe regional disparities in sentiment, such as more nuanced reactions to city-specific policies and spatial spillovers. Our analysis further shows that these GenAI-driven sentiment indices serve as a leading indicator for property sales, and can offer a better proxy for public sentiment compared to news-based sentiment indices.

27. In terms of policy implications, the GenAI-driven sentiment indices should help policymakers better identify trends, understand complex developments and detect and respond to emerging risks. Given the importance of the real estate sector to Mainland China's economy, tracking the recovery of the housing market across regions and in a timely manner is beneficial for both policymakers and market participants. Central banks and other institutions can also use the indices as inputs to nowcast property sales and identify potential turning points.

28. From a methodological perspective, our empirical results highlight the transformative potential of GenAIs in research and macroeconomic surveillance. Firstly, GenAI can make reliable assessments of sentiment as humans, demonstrating its strong comprehension and analytical abilities. Secondly, GenAI can accurately identify "hidden" city names in microblogs among a large volume of data, showcasing its strong reasoning skills and background knowledge. By combining these features with their efficiency, we can now tackle tasks that were previously deemed infeasible, establishing GenAI as a valuable and promising option for applications in big and unstructured data.

REFERENCES

Affuso, E., & Lahtinen, K. D. (2019), "Social media sentiment and market behavior." *Empirical Economics*, 57, 105-127.

Bank of International Settlement (2024a), "Artificial intelligence in central banking.", *BIS Bulletin*, No 84.

Bank of International Settlement (2024b), "Artificial intelligence and the economy: Implications for central banks." *BIS Annual Economic Report*.

Chen, K., & Zhao Y. (2024), "Chinese housing market sentiment index: A generative AI approach and an application to monetary policy transmission." *IMF Working paper, 264*.

Ding, Y., Lee, C., & Lu, M. (2023), "Does market sentiment push up China's housing prices? An empirical study based on the data of 45 mainstream cities in China." *Journal of Housing and the Built Environment*, 38(2), 1119-1147.

Gorodnichenko, Y., Pham, T., & Talavera, O. (2021), "Social media, sentiment and public opinions: Evidence from# Brexit and# USElection." *European Economic Review*, 136, 103772.

Hansen, A. L., & Kazinnik, S. (2023), "Can ChatGPT decipher Fedspeak?" Available at SSRN.

Li, J., Wang, Y., & Liu, C. (2022), "Spatial effect of market sentiment on housing price: Evidence from social media data in China." *International Journal of Strategic Property Management*, 26(1), 72-85.

Liu, B. (2010), "Sentiment analysis and subjectivity." *Handbook of Natural Language Processing*, 2, 627–666.

Lopez-Lira, A., & Tang, Y. (2023), "Can ChatGPT forecast stock price movements? Return predictability and large language models." arXiv preprint arXiv:2304.07619.

Loughran, T., & McDonald, B. (2011), “When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks.” *The Journal of Finance*, 66(1), 35-65.

Ruan, Q., Wang, Z., Zhou, Y., & Lv, D. (2020), “A new investor sentiment indicator (ISI) based on artificial intelligence: A powerful return predictor in China.” *Economic Modelling*, 88, 47-58.

Shao, J., Hong, J., Wang, X., & Yan, X. (2023), “The relationship between social media sentiment and house prices in China: Evidence from text mining and wavelet analysis.” *Finance Research Letters*, 57, 104212.

Shapiro, A. H., Sudhof, M., & Wilson, D. J. (2022), “Measuring news sentiment.” *Journal of Econometrics*, 228(2), 221-243.

Tetlock, P. C. (2007), “Giving content to investor sentiment: The role of media in the stock market.” *The Journal of Finance*, 62(3), 1139-1168.

Zhou, Z. (2018), “Housing market sentiment and intervention effectiveness: Evidence from China.” *Emerging Markets Review*, 35, 91-110.

Zhu, E., Wu, J., Liu, H., & Li, K. (2023), “A sentiment index of the housing market in China: Text mining of narratives on social media.” *The Journal of Real Estate Finance and Economics*, 66(1), 77-118.

Annex A

Topics used for Weibo identification and web scraping (in Chinese)

Topic			
#房地產#	#樓市#	#房產#	#開發商#
#房企#	#房價#	#買房#	#賣房#
#房貸#	#商品房#	#保障房#	#首付#
#公積金#	#新房#	#一手房#	#二手房#

Super topic		
#房地產超話#	#樓市超話#	#房價超話#

Topics used for Weibo identification and web scraping (English translation)

Topic		
#RealEstate	#HousingMarket	#Property
#PropertyDeveloper	#RealEstateCompany	#HousingPrice
#HomeBuying	#HomeSelling	#Mortgage
#CommercialHousing	#AffordableHousing	#HousingProvidentFund
#NewHomes	#FirstHandProperty	#SecondHandProperty

Super topic		
#RealEstate	#HousingMarket	#HousingPrice

Annex B

Examples of microblog texts that are excluded from our sample (in Chinese)

Date	Content	Reason
07 Jan 2024	2023 年悉尼房市回顧：房價驚人回彈！超過四十個郊區房價中位數突破百萬大關任何上升的東西都有下降的時候，反之亦然。2023 年，悉尼的房地產市場證明了這一點，因為悉尼的房屋價值打破了人們的預測，穩步上升，在某些地區甚至上漲了 150 萬澳元。#澳洲##房價##上漲#	Unrelated to Mainland China
06 Jan 2017	【新加坡房價連跌三年：2009 年調控至 2013 年見頂後回落】新加坡政府部門 1 月 3 日公布的最新數據顯示，在政府嚴控政策及經濟增長不景氣的背景下，該國 2016 年第四季度房價環比下跌 0.4%，為連續 13 個季度下跌，創 1975 年公佈數據以來最長下跌記錄，2016 年全年房價下跌 3%。#房價#	Unrelated to Mainland China
30 Jul 2024	牙科醫院，口腔醫院，搞分期的，賣房的，別打我電話，煩不煩啊？AI 一樣，問他哪個單位説不出，開口就是一頓吧啦吧啦#電話詐騙##賣房##牙科#	Unrelated to housing market
25 Feb 2021	晚上看了一眼基金，居然沒綠，這幾個新入手的基金我準備三月底再看，就喜歡大漲大跌的基金，今天還進了一點#房地產#等三月底調倉跌到位的白酒。	Unrelated to housing market
15 Jan 2024	朝陽區現房！央企打造！新房精裝修！目前只剩兩套中疊，143 平米 1100 萬不到，付款好價格可以聊！#房地產#	Advertisement
18 Feb 2023	順德中心！靠近廣州！現房！現房！月供 2500，首付僅需 3 萬即可入住！廣州過來 30 分鐘，佛山新城過來 15 分鐘。看房林經理：18620974872 可安排專車接送看房！好房推薦超話#好房推薦##樓市#	Advertisement
05 Mar 2021	#二手房#2021 年 3 月 4 日，廈門二手房成交 193 套，成交面積 18504 m^2；其中二手住宅成交 154 套，成交面積 16004 m^2。	Data reporting

14 Apr 2022	1—5 月份全國房地產開發投資同比下降 4.0%6 月 15 日，國家統計局數據顯示，1—5 月份，全國房地產開發投資 52134 億元，同比下降 4.0%；其中，住宅投資 39521 億元，下降 3.0%。#房地產#	Data reporting

Examples of microblog texts that are excluded from our sample (English translation)

Date	Content	Reason
07 Jan 2024	2023 Sydney Housing Market Review: Home Prices Surge Back! Over forty suburbs have seen their median home prices surpass the one million mark. Everything that rises must eventually fall, and vice versa. In 2023, Sydney's real estate market proved this point, as home values defied expectations and steadily increased, with some areas even rising by AUD 1.5 million. #Australia #HousingPrices #Increase	Unrelated to Mainland China
06 Jan 2017	[*Singapore's housing prices have fallen for three consecutive years: from 2009 to 2013, they peaked and then fell back*] The latest data released by the Singapore government department on January 3 shows that under the background of the government's strict control policies and sluggish economic growth, in the fourth quarter of 2016, house prices in Singapore fell by 0.4% from the previous quarter, marking the 13th consecutive quarter of decline and the longest decline since data was released in 1975. House prices fell by 3% for the whole of 2016. #Price	Unrelated to Mainland China
30 Jul 2024	Dental hospitals, oral clinics, those offering installment plans, real estate sellers—please don't call me! Isn't it annoying? Just like AI, when you ask it which organization, it can't say anything specific and just rambles on. #PhoneScams #RealEstate #Dentistry	Unrelated to housing market
25 Feb 2021	Took a look at the funds at night, and they didn't go down. I plan to look at these newly purchased funds at the end of March again. I like funds that have experienced sharp rises and falls. I also bought a little #real estate stocks today and will adjust my positions at the end of March when the white wine stocks fall.	Unrelated to housing market
15 Jan 2024	Completed houses in Chaoyang district! Developed by a state-owned enterprise! Newly renovated homes! Currently, only two duplex units are left, 143 square	Advertisement

	meters for under RMB 11 million. Flexible payment options—let's discuss the price! #Real Estate	
18 Feb 2023	Shunde Center! Close to Guangzhou! Completed house! Completed house! The monthly payment is RMB 2,500, and the down payment is only RMB 30,000 to move in! It takes 30 minutes to get there from Guangzhou and 15 minutes to get to Foshan New Town. Please contact Manager Lin: 18620974872, can arrange a special car to pick up and drop off to see the house! #GoodHouseRecommendation #Propertymarket	Advertisement
05 Mar 2021	#SecondHandHouse On March 4, 2021, 193 second-hand houses were sold in Xiamen, with a transaction area of 18,504 square meters. Among them, 154 were residential second-hand houses, with a transaction area of 16,004 square meters.	Data reporting
14 Apr 2022	National real estate development investment fell by 4.0% year-on-year from January to May. On June 15, data from the National Bureau of Statistics showed that from January to May, national real estate development investment was 5.2134 billion yuan, a year-on-year decrease of 4.0%; of which, residential investment was 3.9521 billion yuan, a decrease of 4.0% year-on-year. 3.0%. #RealEstate	Data reporting

Annex C

Examples of microblog scores by in-house GenAI (in Chinese)

Date	Content	Score
19 May 2024	此次政策屬於中國購房歷史上最寬鬆的首付政策，非常重磅，對於拉動市場交易具有非常明顯的作用。過去全國購房首付比例最低為 20%，此次政策市場反饋非常強烈，和過去的限購放鬆、認房不認貸相比，首付比例的下調，其能級遠遠超過其他城市的政策，甚至可以理解為歷史上最寬鬆的一條政策。說明國家層面對於去庫存和支持合理住房消費需求高度重視，其對於房貸方面的拉動、對於剛需和改善型住房的快速釋放和規模性釋放等都具有非常重要的意義。該政策對於房地產市場和地產股等會產生非常重量級的積極影響。#房地產#	9.5 (Positive)
12 Dec 2020	我為什麼不信房價會跌？大城市的就業崗位多，工資高；營商環境好；獲得的社會資源多；能掙更多到錢；身邊都是差不多的人，大家比較有共同語言；意識形態也差不多；治安環境也更好；集中了最好的教育和醫療。單凡是個對生活有追求的人都不會待在鶴崗、小縣城這種地方人都往大城市跑了你說京滬廣深杭會跌嗎？通貨膨脹、城鎮化人口聚焦，每個人的腰包一年比一年鼓，怎麼就有人會覺得房價會跌#買房##深圳樓市##杭州樓市#	9.5 (Positive)
08 Jul 2020	在疫情影響和複雜的環境下，中國樓市的趨勢，有着更為重要的目的。政策和宏觀對樓市的持續管控，只是為了通過橫盤緩衝資本湧入，控制市場盲目信心，以期消化風險，而非容忍房價下跌。要知道，這兩年，無論經濟形勢如何變幻，“穩樓市”都是不變的基調，既要遏制大漲，又要防範大跌，這是底線所在。#樓市##買房##房產中介#	5 (Neutral)
20 Feb 2024	2024 年房價是漲是跌？目前市場上有兩種聲音。第一種聲音是最多的，也就是認為房價要大跌的，理由主要有：1、人口出現負增長，房地產需要人口支撐；2、目前存量房過剩，供需關係發生轉變，過剩必將導致房價降低；3、目前經濟大環境不好，錢難掙，大家不願負債買房。第二種聲音是房價即將觸底反彈，目前是買房的好時機，理由主要有：1、國家多次強調房地產仍是是支柱產業，房住不炒為的是穩定房價避免房價上漲過快，重點是穩而不是想要暴	5 (Neutral)

	跌，隨着後期經濟好轉房價還是會平穩上漲；2、近年來國家連續對樓市的調整能夠看出政府樓市的態度；3、美國加息即將結束，後期外資進場極易導致通貨膨脹，房產仍然是最具有保值的產品之一；4、對於商品房過剩其實只要再次開啓（目前也已經開始了）城中村改造或者拆除會再次出現大量需求。以上是兩種聲音主要依據的理由。大家認為 2024 年房價會漲還是跌呢？#中國人口開啓負增長##未來房價走勢如何##房地產##樓市雜談##房產##買房#	
19 Mar 2024	#房地產#隨便聊兩句。我覺得房地產市場很難再現往日輝煌。再多的刺激政策，都需要靠需求支撐，即得有人購買才可以。但目前的情況是，我國的城鎮化接近尾聲，能進城的年輕人，可能基本都已經進城了。增量用户比較有限，存量用户的需求主要集中在換房，但這個需求量肯定比不上在過去 20 年多年間，大量農村人進城購房的需求量。簡單來講，我覺得一個時代基本結束了。未來即便各個城市間互相逐漸放開限購限貸等各種政策，主要也是對存量用户的爭奪。比如，假設北京放開限購等政策後，對住房的需求開始上漲，則北京周邊大概率會下降。存量用户就這麼多，往這邊跑了，則那邊就少了。	2.5 (Negative)
03 Dec 2021	現在房地產市場的預期已經完全改變：因為房地產的流動性已經枯竭了，而且越來越枯竭。現在千萬不要有投資房地產的想法了，因為源源不斷的，新盤在供應。過去 40 年獲利的二手房都要獲利出局，這個天量的出貨需求。但現在接盤俠在哪裏？投資需求一旦消失，接盤俠連承接市場40%的購買力都沒有。所以説再次強調有多套房的趕緊出逃吧，抓緊打折，甩掉。後面房地產越往後你越賣不動，因為時間橫的越久，會有越多拋售盤。現在昂貴的學區房，包括各個區位的所謂高價好房子，你們要考慮的接盤俠在哪裏？接盤俠已經消失了！現在連資產配置我都已經不建議了，因為我們的房價已經遠遠的透支了未來十年，20 年的潛力，那基本買了就是不動產，一動不動。#房地產#	0.5 (Negative)

Examples of microblog scores by in-house GenAI (English translation)

Date	Content	Score
19 May 2024	This policy represents one of the most lenient down payment policies in the history of home buying in China. It is very significant and has a clear impact on stimulating market transactions. Previously, the minimum down payment ratio for home purchases nationwide was 20%. The market response to this policy has been very strong. Compared to past relaxations of purchase restrictions and the distinction between property recognition and loan recognition, the reduction in the down payment ratio is far more impactful than policies in other cities and can even be understood as the most lenient policy in history. This indicates that the national level places great importance on reducing inventory and supporting reasonable housing consumption demands. It has significant implications for stimulating mortgage lending and the rapid and large-scale release of demand for both first-time buyers and those seeking improved housing. This policy will have a very substantial positive impact on the real estate market and real estate stocks. #RealEstate	9.5 (Positive)
12 Dec 2020	Why don't I believe that house prices will fall? Big cities have more jobs, higher wages, a better business environment, more social resources, more money to earn, surrounded by similar people, we have a common language, similar ideology, a better security environment, and a concentration of the best education and health care. Any person who pursues life will not stay in places like Hegang or small counties and will go to big cities. Do you think the housing prices in Beijing, Shanghai, Guangzhou, Shenzhen and Hangzhou will fall? Given inflation, urbanization, and more people getting more wealthy, how will someone feel that the price of housing will fall? #BuyingHouse #ShenzhenPropertyMarket #HangzhouPropertyMarket	9.5 (Positive)
08 Jul 2020	Under the influence of the pandemic and a complex environment, the trend of China's real estate market has a more significant purpose. The continuous regulation of the real estate market through policies and macroeconomic measures is aimed at buffering capital inflows and controlling blind market confidence	5 (Neutral)

	in order to digest risks, rather than tolerating a decline in housing prices. It is important to note that in the past two years, regardless of how the economic situation has changed, the tone "stabilizing the real estate market" has not changed. The goal is to curb significant price increases while also preventing drastic declines; this is the bottom line. #RealEstateMarket #BuyingHouse #RealEstateAgency	
20 Feb 2024	Will housing prices rise or fall in 2024? Currently, there are two main opinions in the market. The first and most prevalent opinion is that housing prices will drop significantly, with the main reasons being: 1. There is negative population growth, and real estate needs population support; 2. There is currently an oversupply of existing homes, leading to a shift in the supply-demand relationship, and this oversupply will inevitably result in lower prices; 3. The overall economic environment is poor, making it difficult to earn money, and people are reluctant to take on debt to buy homes. The second opinion is that housing prices are about to hit bottom and rebound, suggesting that now is a good time to buy. The main reasons for this view are: 1. The government has repeatedly emphasized that real estate remains a pillar industry, and the principle of "housing for living, not for speculation" aims to stabilize prices and prevent rapid increases, focusing on stability rather than a desire for a crash. As the economy improves, housing prices are expected to rise steadily; 2. In recent years, the government's continuous adjustments to the housing market indicate its stance; 3. The end of interest rate hikes in the U.S. may lead to an influx of foreign capital, which could result in inflation, making real estate one of the most valuable assets; 4. Regarding the oversupply of commercial housing, as long as urban village renovations or demolitions are restarted (which has already begun), there will be a resurgence in demand. These are the main reasons behind the two opinions. What do you think will happen to housing prices in 2024? #ChinaPopulationNegativeGrowth #FutureHousingTrends #RealEstate #HousingMarketTalk #Property #BuyingHouse	5 (Neutral)

19 Mar 2024	#RealEstate Just a few thoughts. I believe the real estate market is unlikely to regain its past glory. No matter how many stimulus policies are introduced, they all need to be supported by demand—there must be buyers. However, the current situation is that urbanization in our country is nearing its end, and most of the young people who can move to cities have already done so. The number of new users is quite limited, and the demand from existing users mainly revolves around upgrading their homes. But this demand is certainly not comparable to the massive influx of rural residents buying homes in the past 20 years. In simple terms, I think an era has essentially come to an end. In the future, even if cities gradually relax purchase and loan restrictions, it will mainly be a competition for existing users. For example, if Beijing lifts its purchase restrictions and housing demand starts to rise, it is likely that demand in surrounding areas will decrease. There are only so many existing users; if they move in one direction, there will be fewer in the other.	2.5 (Negative)
03 Dec 2021	Now the real estate market expectations have completely changed: because the liquidity in real estate has dried up, and is getting dryer. Don't even think about investing in real estate now, because there is a constant supply of new properties. Second-hand houses that have made profits in the past 40 years will all be sold out due to the huge demand for shipments. But where is the buyer now? Once the investment demand disappears, the buyer will not even have 40% of the purchasing power in the market. So I emphasize once again that if you have multiple apartments, grab the discount and get rid of it. The longer the time goes by, the less likely you will be able to sell it because the more selling orders there will be. Nowadays, for houses in expensive school districts, including so-called high-priced and good houses in various locations, where are the buyers? The buyers have disappeared! Now I don't even recommend asset allocation, because our housing prices have far exceeded the potential of the next ten or 20 years, so what we buy is literally "real estate", which will not change. #RealEstate	0.5 (Negative)

Annex D

Examples of city identified by in-house GenAI (in Chinese)

Date	Content	City
02 Jan 2020	#房價#降準 0.5 對樓市根本沒啥影響，一些狗屁大 V 就開始忽悠房價要漲，純粹滿嘴噴糞，胡説八道！給他們自己解套忽悠接盤俠呢！2020 帝都房價會繼續跌跌不休，現在大體上跌到了 2016 下半年的水平，2020 將跌回 2016 上半年，至於會跌到 2016 年上半年的幾月，拭目以待！	北京
22 Jul 2020	首創和遠洋下午拿了豐台大瓦窯兩塊地，現在就有銷售喊着這兩地能賣 8+10+的。別聽忽悠！地塊旁邊的萬科假日風景 6+二手房掛半年還沒賣出去呢，開發商可能自己都沒想明白這地賣給誰#房地產#	北京
12 Jun 2022	唐山房價要大跌了，唐山房價現在 1.2 萬一平米在河北省排名第 2，經過打人事件的負面影響唐山房價要跌到河北省第 5，最悲觀的情況，可能要跌到第 7 了。來看一下河北省 11 個城市的房價情況 1，石家莊，房價 1.46 萬一平米 11，張家口，房價 0.71 萬一平米，石家莊的房價是張家口的 2 倍石家莊是河北省會，但整個河北省，離北京，天津太近石家莊這個省會效應不大大量河北人口，都去北京，天津發展了。2，唐山，1.2 萬一平米 3，廊坊市，1.18 萬一平米 4，秦皇島，1.17 萬一平米 5，滄州市，1.15 萬一平米這四個城市，房價差距很少唐山，單單投資客的撤離，房價至少要跌幾百鐵定要跌到第五了。6，保定市，1.09 萬一平米 7，邯鄲市，1.03 萬一平米與唐山的房價只差一千多，唐山這次事件，後續的發展，影響力還要繼續觀察，如果後續繼續爆出唐山的各種負面消息的話，那房價將進入持續下跌的悲慘境地。8，承德市，房價 0.91 萬一平米 9，衡水市，房價 0.77 萬一平米 10，邢台市，房價 0.76 萬一平米。#房地產#	唐山
29 Jun 2022	#樓市#魔都 6 月豪宅成交情況，和天氣一樣🔥	上海
03 Sep 2023	朋友圈一張圖傳的飛起！上海新房漲了 12%我冒昧問一下上海上漲和武漢有什麼關係？反正我覺得沒有啥關係，別人上海的是超一線城市是魔都，跟武漢的級別還是有差距而且武漢的樓市跟上海的更是天差地別！這次的認房不認貸對樓市是非常利好嗎？不見得吧，不就是把老韭菜拿出來再割一割	武漢

	麼？還憋了那麼久，沒必要，這種政策真沒必要！現在是要解決大家賺錢的問題！賺不到錢你再怎麼省錢有毛用，每個月少那 300、500 的。又不是一次性在首付上總價上少幾十萬有啥？信心早早的都被玩沒了，重鑄信心才能改變目前地產格局！#武漢##房地產##經濟#	

Examples of city identified by in-house GenAI (English translation)

Date	Content	City
02 Jan 2020	#HousingPrices The reserve requirement ratio cut of 0.5 has no real impact on the real estate market. Some so-called influencers are starting to hype that housing prices will rise, which is just pure nonsense and ridiculous talk! They're just trying to save themselves by misleading new investors! In 2020, housing prices in the Imperial City will continue to decline without end. They have generally dropped to the levels of the second half of 2016, and in 2020, they will fall back to the first half of 2016. As for how far back into the first half of 2016 they will drop, we'll have to wait and see!	Beijing
22 Jul 2020	Beijing Capital Land and Sino Ocean acquired two plots of land in Fengtai's Dawayao area this afternoon, and now some salespeople are claiming that these plots can be sold for 8+10+ (note: the numbers were likely referring to a price range of RMB 80k to 100k per square meter). Don't believe the hype! The Wanke Holiday Scenic Area, which is right next to these plots, has had a 6+ (likely referring to a price range of RMB 60k per square meter) resale listing for half a year and still hasn't sold. The developers themselves might not even have a clear idea of who they can sell this land to #RealEstate	Beijing
12 Jun 2022	1. Tangshan's housing prices are going to plummet! Currently, Tangshan's housing prices are at RMB 12,000 per square meter, ranking 2nd in Hebei Province. However, due to the negative impact of the recent violent incident, Tangshan's housing prices are expected to drop to 5th place in the province. In the worst-case scenario, prices might even drop to 7th place. Let's take a look at the housing prices in the 11 cities of Hebei Province: Shijiazhuang: RMB 14,600 per square meter, Zhangjiakou: RMB 7,100 per square meter, Shijiazhuang's housing prices are twice that of	Tangshan

	Zhangjiakou. As the provincial capital, Shijiazhuang's prices should be higher, but since Hebei Province is too close to Beijing and Tianjin, the provincial capital effect is not significant. Many people from Hebei have gone to Beijing and Tianjin for development. 2. Tangshan: 12,000 RMB per square meter. 3. Langfang: 11,800 RMB per square meter. 4. Qinhuangdao: 11,700 RMB per square meter. 5. Cangzhou: 11,500 RMB per square meter. These four cities have very similar housing prices. With the withdrawal of investors from Tangshan, housing prices are expected to drop by at least a few hundred RMB, and will definitely drop to 5th place. 6. Baoding: 10,900 RMB per square meter. 7. Handan: 10,300 RMB per square meter. The price difference between these cities and Tangshan is only a few hundred RMB. The impact of the recent incident in Tangshan is still being observed, and if more negative news about Tangshan comes out, the housing prices will enter a state of continuous decline. 8. Chengde: 9,100 RMB per square meter. 9. Hengshui: 7,700 RMB per square meter. 10. Xingtai: 7,600 RMB per square meter. #RealEstate	
29 Jun 2022	#RealEstate The luxury home transaction situation in the Magic City in June is just like the hot weather.	Shanghai
03 Sep 2023	The picture shared in the circle of friends is going viral! New home prices in Shanghai have risen by 12%. May I ask what the relationship is between the rise in Shanghai and Wuhan? Anyway, I don't think there's much of a connection. Shanghai is a super first-tier city, the 'Magic City,' and there's still a gap in status compared to Wuhan. The real estate markets in Wuhan and Shanghai are worlds apart! Is this new policy of recognizing homes but not loans really beneficial for the real estate market? I don't think so; it's just about taking advantage of the old investors again, right? After holding back for so long, it’s unnecessary. This kind of policy is really not needed! What we need to do now is solve the problem of people making money! If you can't make	Wuhan

	money, what's the use of saving a little here and there, like RMB 300 or RMB 500 each month? It's not like saving tens of thousands on the total price of the down payment all at once means anything. Confidence has already been lost early on; we need to rebuild confidence to change the current real estate landscape! #Wuhan #RealEstate #Economy	

Some explanation is as follows.

- In the examples dated "02 Jan 2020" and "29 Jun 2022", the city names are not explicitly stated; yet, our GenAI accurately infers "Beijing" from the city's nickname on social media, "Imperial City" and "Shanghai" from "Magic City", respectively. This demonstrates the system's ability to identify the city under focus even when the city name is not directly mentioned.

- In the second example dated "22 Jul 2020", the city name "Beijing" is not present in the content; yet, our GenAI is still able to identify it as the city under focus. This is possible because the post mentions "Fengtai", a district within Beijing, which our GenAI recognises as a valid reference to the city. This highlights the system's ability to identify a city even when only a district within that city is mentioned.

- In the third example dated "12 Jun 2022", multiple cities are mentioned, but the GenAI correctly identify the city being discussed. Similarly, the last example dated "03 Sep 2023" presents a scenario in which two cities, Shanghai and Wuhan, are explicitly mentioned in the post. However, our GenAI correctly identifies Wuhan as the city under focus. This demonstrates its capacity to distinguish between relevant and comparative information.