Chinese Housing Market Sentiment Index: A Generative AI Approach and An Application to Monetary Policy Transmission

Kaiji Chen (Emory University), Junru Guo (Nankai University), Jia He (Nankai University), and Yunhui Zhao (IMF)[1]

September 2024

ABSTRACT: We construct a daily Chinese Housing Market Sentiment Index by applying GPT-4o to Chinese news articles. Our method outperforms traditional models and some other generative AI models in several validation tests, including a test based on a suite of machine learning models. Applying this index to household-level data, we find that after monetary *easing*, homebuyers in cities with more optimistic housing sentiment have *lower* responses in non-housing consumption due to the larger increase in house prices in these cities. Methodologically, our paper offers a tool for timely monitoring housing sentiment and lays out some principles for applying generative AI models, adaptable to other studies globally.

JEL Classification Numbers:	C10, C53, E31, E37, R30
Keywords:	Chinese Housing Market Sentiment, Generative AI, Monetary Policy Transmission, Consumption, Crowding-Out
Authors' E-Mail Addresses:	kaiji.Chen@emory.edu; 1120200947@mail.nankai.edu.cn; HeJia@nankai.edu.cn; YZhao@imf.org

[1] We thank Jin Zhang (Emory University) for her excellent research assistance. We are also grateful for the helpful discussions with Mehdi Benatiya Andaloussi (all IMF, unless specified otherwise), Tohid Atashbar, Nina Biljanovska, Alessia De Stefani, Chris Evans, Nan Li, Li Liu, Rui Mano, Jorge Leon Murillo, Lev Ratnovski, Yannick Timmer (Federal Reserve Board), Stijn Van Nieuwerburgh (Columbia University), Jiaxiong Yao, and participants in the Research Ideas Workshop at the IMF in June 2023 and 2024 ABFER-JFDS Conference on AI and Fintech. He thanks the National Natural Science Foundation of China (Nos. 72271130 and 71901128) for financial support. The views expressed herein are those of the authors and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

I. INTRODUCTION

Monitoring housing market sentiment is crucial for households, financial institutions, and policymakers in any country with a large housing sector. This is especially the case at the current conjuncture for China, which is suffering from the housing distress following the 2021 Evergrande liquidity crisis. In response, the Chinese authorities have recently eased monetary policy stance to boost domestic consumption and stimulate the economy. These raise critical questions: How can we accurately measure housing market sentiment in China? And how does this sentiment affect monetary policy transmission and its impact on consumption in China?

News articles provide a natural opportunity for gauging market sentiment, but traditional methods face challenges in properly capturing the nuanced sentiments expressed in these articles. Artificial intelligence (AI), particularly generative AI-powered Large Language Models (LLMs) such as GPT-4o, holds great potential in addressing these challenges due to its advanced capabilities in processing and understanding complex text data. Unlike keyword-based models, which often miss context and nuance, GPT-4o utilizes deep learning to understand the context and subtleties in language. And compared to some other LLMs such as BERT, which require extensive labeled data for training, GPT-4o can perform zero-shot and few-shot learning,[2] making it more flexible for analyzing diverse and dynamic text data.

In this paper, we apply OpenAI's GPT-4o model to construct a Chinese housing market sentiment index (CHMSI) using news article data. We present multiple versions of the index, each created with different levels of "prompts"—instructions provided to the generative AI models. We summarize eight effective principles for "prompt engineering," a process involving designing and refining input instructions to guide AI models. We find that the housing sentiment index constructed by GPT-4o with "advanced" prompts outperforms those constructed by the traditional keywords-based model and many Chinese LLMs according to all three criteria: First, consistency with human assessments. Using selected testing articles, the housing market sentiment scores generated from GPT-4o model with advanced prompts exhibits the lowest sum of square errors relative to human assessment. Second, the ability to capture market sentiment around major housing policy announcements. Compared with their counterparts generated from other LLMs, the housing sentiment index constructed by

[2] Zero-shot prompting refers to the ability of a model to perform a task without having seen any examples of the task during training, relying instead on the contextual understanding of the prompt alone. Few-shot prompting involves providing the model with a small number of examples (usually one to five) to help it understand the task better before performing it on new data.

GPT-4o exhibits less noise and more conformity with announcements in major housing policies. Third, the ability to predict national housing price trends with machine learning models. We apply different sentiment indices to machine-learning models, commonly employed in macroeconomic forecasting, to predict the year-on-year growth rate of house prices and perform a "horse race" of the models to determine the best performing forecast models. Our optimal forecasting model (the random forest model) using GPT-4o constructed sentiment index achieves a higher forecasting performance than those using the traditional keywords-based approach or the Chinese LLMs.

After establishing the validity of the sentiment index constructed by GPT-4o according to the above-mentioned three criteria, we extend our daily sentiment index data to September 11, 2024. Doing so enables us to trace the housing sentiment during the ongoing housing slump in China as well as following recent major polices. Our up-to-date Chinese housing market sentiment index can well trace the housing sentiment around important changes in housing regulatory policies (such as the "three-red-line" policy) and during housing market distresses (such as those associated with the financial difficulties of Evergrande and Country Garden), and it can shed light on the impact of Chinese government housing stimulus policies on household confidence in housing markets.

We then apply our CHMSI to evaluate the role of the housing sector and the effectiveness of China's monetary easing in stimulating domestic consumption. We employ a proprietary household-level dataset from a leading Chinese commercial bank between 2012Q1 and 2016Q1, covering transaction-level panel data on consumption of various categories and transaction-level repeated cross-sectional data on mortgage originations. Our monetary policy shocks are from Chen, Ren, and Zha (2018), constructed using an endogenous switching M2-based monetary policy rule. Our empirical strategy exploits variations across Chinese cities in exposures (to the national-level housing market sentiment), measured by the city-level Baidu search index, which serves as a proxy for the attention of a city's potential homebuyers and for other shocks to the city's local housing markets. Specifically, by incorporating Baidu search index normalized by city-level population, we refine the CHMSI to create an "Attention-Adjusted Chinese Housing Market Sentiment Index" (AACHMSI) at the city-level. We classify cities according to AACHMSI across all cities and quarters and compare the responses of non-housing consumption to monetary shocks between households in cities belonging to the top 20th percentile in terms of AACHMSI (i.e., being in an "optimistic regime") and those in cities belonging to the bottom 20th percentile (i.e., being in a "pessimistic regime"). The difference between the average responses of non-housing consumption to monetary policy shocks for households in

these two groups of cities provides a creditable estimate of the effects of housing market sentiments on monetary policy transmission into non-housing consumption.

We use a local projection approach to estimate the interactive effects of housing market sentiments on the monetary policy transmission into non-housing consumption. Our local projection results suggest that following monetary easing, the non-housing consumption of homebuyers in cities under *optimistic* housing sentiment increases by *less*, particularly for households with a college degree and aged between 30 and 50. Moreover, we find that following monetary easing, house prices increase more in cities under optimistic housing sentiment than those under non-optimistic housing sentiment, suggesting that the higher house prices may be a channel through which the aforementioned crowding-out occurs, that is: following monetary easing, house prices become higher, which increase down payments and future debt-servicing costs and crowd out non-housing consumption.[3] We interpret these findings by developing a simple analytical model that features the "trade-up" of houses by existing homeowners, i.e., upgrading to larger houses.

Our paper contributes to four strands of literature. First, it contributes to the emerging literature on applying machine learning and LLMs to research in household finance, especially for topics related to mortgage and housing decisions. For example, Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2022) evaluate the change from traditional logit technology to machine learning technologies using a large dataset from the U.S. mortgage market. They find that Black and Hispanic mortgage borrowers in the U.S. are disproportionately less likely to benefit from the introduction of machine learning into default prediction models (see also Sadhwani, Giesecke, and Sirignano, 2021, and Barbaglia, Manzan, and Tosetti, 2023 for the application of machine learning algorithms in predicting mortgage loan defaults).[4] Apart from machine learning approaches, researchers have been increasingly using LLMs in the field of finance, especially in predicting stock returns. However, to the best of our knowledge, this paper is one of the first to apply LLMs in household finance and especially in studies on

[3] These findings suggest that monetary easing might be more effective in boosting household consumption at the current conjuncture than in the past for China, as the crowding-out channel may be weaker under the prevailing pessimistic housing sentiment. Moreover, for monetary policy to be effective in boosting household consumption at the current conjuncture, it is important for other policies, such as housing policies and structural reforms, to coordinate with monetary easing to contain housing speculation.

[4] See also Albanesi and Vamossy (2019), Medeiros et al. (2021), Korinek (2023), Bartik, Gupta, and Milo (2023), among others, on the application of machine learning tools and LLMs to other areas in economics and finance.

housing and monetary policy transmission. In terms of research questions, rather than using LLMs to predict asset prices or returns, our goal is to use LLMs to analyze unstructured texts from news media to construct housing sentiments, a crucial factor governing households' expectation of house price growth. Although there is growing consensus on the importance of housing market speculation or optimism in the U.S. housing booms and busts of 2000 to 2010, there is limited individual-level data on house price beliefs during the housing boom from 2002 to 2006 that can be linked to measures of speculation at the individual level (Mian and Sufi, 2022).[5] Hence, our approach of using GPT-4o to construct disaggregated housing market sentiments with news article data can be applied to other countries, such as the U.S., thus providing a useful complement to the literature.

Second, our paper contributes to the extensive literature on the effects of monetary policy on household consumption using granular data.[6] For example, Agarwal et al. (2022) show that transmission from mortgage rates to households' non-housing consumption in China is present only for homeowners with unpaid mortgage balances and not for other types of households.[7] Many of these studies use repeated cross-sectional data (e.g., Cloyne, Ferreira, and Surico, 2020) or impute consumption using administrative data (e.g., Holm, Paul, and Tischbirek, 2021). As noted by Baker et al. (2022) and Cao et al. (2024), since imputed consumption is constructed as income not saved, it is not possible to identify what type of goods or services consumers acquire. Our transaction-level credit card data, nonetheless, can distinguish between different types of consumption expenditures, and thus we construct non-housing consumption

[5] After the Global Financial Crisis (GFC), there are household survey data available to measure household beliefs on the housing market. However, no such data is available before then, and our method can be applied to construct similar housing market sentiment before the GFC using, for example, news article data in the U.S.

[6] See, among others, Agarwal, Liu, and Souleles (2007), Agarwal and Qian (2014), Di Maggio et al. (2017), Chen et al. (2020), Agarwal et al. (2022), and Beraldi and Zhao (2023).

[7] Another view emphasizes the wealth channel of monetary policy transmission. For example, Painter, Yang, and Zhong (2022) empirically show that the marginal propensity to consume out of current wealth in China is low. Moreover, as pointed out by the paper, households are prohibited from withdrawing housing equity in China, further limiting the wealth channel. Relatedly, although a large number of studies (such as Jarocinski and Smets (2008), Jordà, Schularick, and Taylor (2015), and Williams (2016)) find that monetary policy shocks have significant effects on house prices (and presumably on consumption), some studies find that this is not always the case: for example, Favilukis, Ludvigson, and Van Nieuwerburgh (2017) find that a large foreign capital inflow to an economy would endogenously increase the housing risk premium, substantially offsetting the effect of lower interest rates on house prices.

by excluding housing-related expenses. With this unique dataset connecting non-housing consumption and mortgage origination at individual levels, we identify a novel crowding-out channel for monetary policy easing to affect non-housing consumption.

Third, our paper contributes to the literature on sentiment analysis in the housing market. Zhou (2018) examines the interaction between housing market sentiment and government intervention in China. The author finds that positive sentiment leads to higher house prices and increased market activity, whereas negative sentiment has the opposite effect. Dong, Hui, and Yi (2021) explore the relationship between housing market sentiment and homeownership in China, concluding that optimistic sentiment boosts homeownership rates by enhancing potential buyers' confidence in future house prices. Cepni and Khorunzhina (2023) construct state-level housing-sentiment indices in the U.S., demonstrating that sentiment significantly influences housing market performance across different states. Furthermore, Cepni (2024) delves into the effects of news media coverage on housing market sentiment across various U.S. states, underscoring the importance of media in molding public perception and market results. Alfano and Guarino (2022) analyze the impact of textual sentiment in news articles on U.K. housing prices, showing that positive news sentiment drives up prices while negative sentiment causes declines. Our paper builds on these studies by using generative AI models to construct a refined and dynamic sentiment index for the Chinese housing market. This index, validated against human assessments, major housing policies, and housing price trends, serves as a robust tool for examining the impact of sentiment on monetary policy transmission and housing consumption.

Finally, our paper contributes methodologically to the literature on how to effectively use generative AI tools. Chen et al. (2023) lay out foundational principles of prompt engineering and discuss advanced methodologies like chain-of-thought prompting.[8] Barbaglia et al. (2023) provide a systematic study of different guiding principles and offer prompting advice for prompting leading LLMs.[9] Our paper extends these principles by incorporating additional effective strategies, such as requesting the code for replication (thereby enhancing transparency). In addition, Sahoo et al. (2024) provide a systematic survey of prompt engineering in LLMs, discussing the advantages and disadvantages of few-shot versus zero-shot prompting. They argue that although few-shot prompting can improve performance for complex tasks, the selection and

[8] Wei et al. (2022) and Zhang et al. (2022) also discuss the Chain-of-Thought prompting and argue that this strategy enhances the performance of LLMs by producing some intermediate reasoning steps.

[9] Relatedly, Korinek (2023) illustrates the automated prompt generator offered by Anthropic PBC, a U.S.-based AI startup.

composition of prompt examples can significantly influence model behavior. Building on these insights, our paper primarily focuses on zero-shot prompting for sentiment analysis. Our contribution to this literature lies in showcasing the use of generative AI in a novel context and providing rigorous validation tests for the generative AI-constructed indices, including tests that use a suite of machine learning models.

The rest of the paper is organized as follows. Section II constructs Chinese Housing Market Sentiment Index (CHMSI). Section III conducts validity tests of the CHMSI we constructed and compares it with indices constructed by other models. Section IV applies our sentiment index to estimate the role of housing market sentiments on the monetary policy transmission into consumption. Section V concludes.

II. CONSTRUCTION OF CHINESE HOUSING MARKET SENTIMENT INDEX

A. Data and Methodology

We utilize daily news article data from the China Stock Market and Accounting Research (CSMAR) news database to construct the Chinese housing market sentiment index.[10] The dataset starts from January 4, 2012 to September 11, 2024, providing a comprehensive up-to-date view of the housing market sentiment from the news media. The CSMAR database collects news articles related to economics and finance published by major Chinese newspapers, which ensures the content is professional and less biased compared to social media data often used in other studies. This high-quality data source enhances the reliability of our sentiment analysis. First, it provides more professional and curated content, reducing the noise often associated with social media. Second, the CSMAR dataset is updated daily, allowing us to construct a highly timely sentiment index. This timeliness is crucial for market participants, researchers, and policymakers who require up-to-date information to make informed decisions. Although we use data from 2012 to 2016 for illustration purposes and to manage the financial costs of running the algorithm, the same methodology can be applied to the most recent data to maintain the index's relevance.

Our methodology involves leveraging the sentiment analysis function in GPT-4o, accessed through the OpenAI API. GPT-4o's advanced capabilities in understanding and processing complex text data make it an ideal tool for this task. It offers significant advantages over other sentiment analysis tools, such as keyword-based models and BERT. While keyword-based models often miss context and nuance, GPT-4o uses deep

[10] The CSMAR database could be accessed via Wharton WRDS database.

learning to capture the subtleties in language. Additionally, GPT-4o supports zero-shot and few-shot learning, making it more flexible and powerful without requiring extensive labeled data.

Specifically, we follow these steps to construct the sentiment index:

- Data retrieving and preprocessing. We retrieve daily news articles from the CSMAR database. The collected articles are preprocessed to clean and standardize the text. This step includes removing irrelevant content, normalizing text (e.g., converting to lowercase, removing punctuation), and tokenizing the text for analysis.
- Sentiment analysis with GPT-4o: Using GPT-4o, we analyze the sentiment of each news article. We design prompts that instruct the AI to accurately assess the sentiment. GPT-4o generates sentiment scores based on the context and nuances captured in the articles, ensuring a sophisticated understanding of market sentiment.
- Prompt engineering. We apply eight effective principles of prompt engineering to enhance the performance of the sentiment analysis, as explained subsequently. Section B lays out the details for our principles of prompt engineering.
- Index construction. We aggregate the daily sentiment scores to construct a comprehensive sentiment index. This index reflects the overall sentiment of the housing market as depicted in the news articles.
- Validation. We validate the constructed sentiment index using three criteria: consistency with human assessments (compare AI-generated sentiment scores with those given by human evaluators), ability to capture market sentiment around major housing policy announcements, and consistency with national housing price evolution.

By following these steps, we ensure that the sentiment index is robust, timely, and reflective of the true market sentiment. The use of GPT-4o for sentiment analysis, combined with effective prompt engineering and rigorous validation, provides a reliable tool for monitoring and analyzing the Chinese housing market.

B. Principles of Prompt Engineering

Given the crucial role of prompt engineering, we lay out eight principals based on the literature and our own experiments. Specifically:

- *Be specific.* Provide clear and detailed instructions in your prompts. Avoid vague or ambiguous language that may lead to misunderstandings. The more precise your prompt, the more accurate and relevant the AI's response will be. For instance, rather than asking, "What do you think about the housing market?", specify, "Analyze the trends in the Chinese housing market over the past five years and predict future developments based on current data."
- *Incorporate role-playing*. Direct the AI to adopt a specific role or perspective to enhance the relevance of its responses. For example, instruct the AI to “Act as an expert in Chinese housing, macroeconomics, and finance, and provide an assessment of the housing market sentiment in China.” This approach helps the AI tailor its response to the context and deliver more insightful analysis.

- *Request replicable code*. When dealing with technical or analytical tasks, request the AI to provide the code or methodology used in its analysis. This ensures transparency and allows for the replication or verification of results. For example, prompt, "Provide the Python code used to calculate the housing market sentiment index based on the given dataset."
- *Iterate and refine.* Employ an iterative approach to refine prompts and improve the quality of the AI's output. Start with a basic prompt and progressively refine it based on the AI's responses. Iteration helps in honing the instructions to achieve the desired level of detail and accuracy. For instance, begin with "Analyze the impact of monetary policy on housing prices" and refine it to "Analyze the impact of recent monetary policy changes on housing prices in major Chinese cities over the past year."
- *Encourage critical evaluation.* Encourage the AI to critically evaluate its responses and consider alternative perspectives. This can be done by including prompts that ask the AI to verify its conclusions or explore different scenarios. For example, prompt, "Assess the impact of rising interest rates on the housing market, and consider if there could be any countervailing factors that might mitigate these effects." In addition, a simple follow-up prompt like "Are you sure?" can often be highly effective.
- *Do not rush.* Explicitly instruct the AI "Don't rush" to allow for sufficient time to process and generate responses. Rushed prompts or expecting immediate answers can lead to suboptimal results. Ensure that the AI has adequate time and context to understand and respond accurately. This might involve setting realistic expectations for response times and avoiding overly urgent prompts.
- *Apply chain-of-thought*. LLMs are capable of performing detailed reasoning processes by creating steps in between, known as chain-of-thought (CoT) prompting. There are primarily two approaches within CoT prompting. The first one encourages sequential thinking with a simple prompt such as "Let's think step by step." The second one involves crafting detailed examples that pair questions with their respective reasoning sequences to arrive at an answer. The effectiveness of this latter method primarily relies on meticulously creating specific demonstrations. These labor-intensive tasks can be reduced by sequentially soliciting demonstrations with a new prompt "let's think not just step by step, but also one by one," as shown by Zhang et al. (2022).
- *Tip and penalize*. Experience suggests that interestingly, strategies such as artificially offering LLMs a tip if they do well or threatening a penalty if not can improve performance, as also documented in Bsharat, Myrzakhan, and Shen (2023).

By adhering to these principles, researchers can enhance the effectiveness of AI-generated responses, thereby achieving more accurate and insightful outcomes in their analyses, as illustrated in the next section. The prompts we use for the rest of the paper also exemplify the implementation of these principles, and the details are provided in Appendix 1 and Appendix 2.

III. VALIDITY TESTS AND COMPARISON WITH OTHER MODELS

This section conducts validity tests of the CHMSI we constructed and compares it with indices constructed by other models. We will do this according to three criteria: human assessments of a few testing sentences, sentiment around announcements of "major housing policies," and forecasting power for house price dynamics based on a suite of machine learning models.

A. Comparison with Human Assessments

Consistent with the literature, comparing with human assessments (or "subjective evaluation") is an essential and intuitive way to gauge the validity of the outputs by LLMs. We first conduct the evaluation based on five sentences we select from the actual news articles in our database: one "benchmark" that has a neutral sentiment based on our manual evaluation, two "positive" sentences that have a clear optimistic tone, and two "negative" sentences that have a clear pessimistic tone. We then conduct the evaluation based on five *articles* selected in a similar way.

Note that although we provide the English versions of these sentences (or titles of news articles) below, we intentionally instruct the LLMs to detect the sentiment directly from the original Chinese text rather than translating first and then detecting the sentiment. Doing so can usually better capture the nuances and subtleties in the original language, as demonstrated by Zhang et al. (2022), among others.

Testing Results Based on Five *Sentences*

The five sentences we select are (in English):[11]

- **Benchmark Sentence:** The earliest development of Luhuitou Peninsula can be traced back to approximately 20 years ago.
- **Positive Sentence 1:** There are concerns in public opinion that continuous relaxation policies will undermine the achievements of many years of real estate market regulation; such concerns are truly unnecessary.

[11] To facilitate replications of our results, we also provide these five sentences in the original language (Chinese): (1) **Benchmark Sentence:** 鹿回头半岛最早的开发可以追溯到大约 20 年前. (2) **Positive Sentence 1:** 有舆论担心，持续的松绑政策会使多年来的楼市调控成果毁于一旦；这种担心实在是不必要. (3) **Positive Sentence 2:** 多家机构预计，监管层很有可能继续放松房地产政策，出台刺激措施遏制房地产业的收缩速度，以缓解低迷的房地产市场. (4) **Negative Sentence 1:** 特别是近期房价上涨过快的热点城市，不排除会尽快出台更加严厉的具体调控政策. (5) **Negative Sentence 2:** 虽然整体而言，此次调控在措施内容上并未超出市场预期，但是从操作力度和中央的决心看，调控有可能再次对地产市场走向和股市形成长期压制.

- **Positive Sentence 2:** Multiple institutions predict that regulatory authorities are likely to continue relaxing real estate policies and introduce stimulus measures to curb the contraction speed of the real estate industry, in order to alleviate the sluggish real estate market.
- **Negative Sentence 1:** Especially in hotspot cities where housing prices have risen too quickly recently, it is not ruled out that more stringent specific regulatory policies will be introduced soon.
- **Negative Sentence 2:** Although overall, the content of this regulation does not exceed market expectations, judging by the operational strength and the central government's determination, the regulation may once again exert long-term pressure on the real estate market and the stock market.

The validity test and comparison results are presented in Table 1, where the sentiment score ranges from 0 (most pessimistic) to 100 (most optimistic). The keyword-based approach significantly deviates from capturing the tone of the broadly neutral benchmark sentence and misinterprets the overall sentiment of both positive and negative sentences. The performance of the Chinese LLMs is mixed: Two such models (Senta and Qianfan) miss the tones of most of the five sentences by a large margin, which could be due to the bias associated with our extremely small sample of five sentences; but the other four (Kimi, Xinghuo, Baihuan, and Hunyuan) produce accurate assessments of the sentiment. As for GPT-4o, all three versions of its sentiment scores (each corresponding to a different level of the prompt) also accurately capture the tone of the benchmark sentence and the sentiment directions of all other sentences.

[Insert Table 1 here]

To more precisely assess the performance of each method and to illustrate the impact of effective prompt engineering, we first normalize the sentiment score by the score assigned to the benchmark sentence. This normalization mitigates concerns regarding differing scaling standards across methods in sentiment score assignment. Subsequently, for each method, we calculate the sum of squared errors (SSE) of its sentiment scores relative to human assessments. The larger the deviation of the overall score assignment from the human assessment of the corresponding sentence (after accounting for scaling differences), the poorer the performance of that sentiment analysis method. The results are presented in Table 2.

[Insert Table 2 here]

Table 2 clearly demonstrates the advantages of some Chinese LLMs and the GPT-4o model (particularly when used with an advanced prompt). The first row of Table 2

presents the normalized results from Table 1, indicating that the "Positive 1" and "Positive 2" sentences are more optimistic than the benchmark sentence (by 40 percent and 60 percent, respectively), and that the "Negative 1" and "Negative 2" sentences are more pessimistic.[12] Consistent with Table 1, the keyword-based method displays a high sum of squared errors (SSE) (2.18), indicating significant deviations from human assessments. Again, the use of two Chinese LLMs (Senta and Qianfan) does not improve the sentiment analysis. However, the use of the other six Chinese LLMs and the GPT-4o model with a "basic prompt" dramatically enhances the sentiment analysis. As the quality of the prompt improves, the SSE of the GPT-4o model decreases monotonically, suggesting a gradual improvement in the quality of the sentiment analysis. Note that the GPT-4o model with the advanced prompt is still outperformed by three Chinese LLMs (Kimi, Xinghuo, and Baichuan), a point we will revisit later.

We now examine *why* the different approaches lead to dramatically different sentiment assessments. Table 3 presents the rationales of the keyword-based approach for its sentiment assessments. This approach, like many traditional sentiment analysis methods, counts the number of positive and negative words based on a predefined dictionary and defines the sentiment score as the (scaled) ratio of the net number of positive words (i.e., the number of positive words minus negative words) over the total number of words. The disadvantage of this method is most evident in the "Positive 2" sentence, where it counts "遏止" (curb) and "收缩" (contraction) as negative words and fails to capture that these words combined in the phrase "curb the contraction" convey a positive sentiment.

[Insert Table 3 here]

By contrast, the GPT-4o models with different levels of prompts can capture nuances, logical connections, and subtleties, reflecting not only the direct market outlook but also the indirect implications of government regulations. These nuances are shown in Table 4 for the basic prompt and Table 5 for the advanced prompt. Comparing Table 5 with Table 4 clearly demonstrates that the GPT-4o model with an advanced prompt provides more detailed, rigorous, and clearer reasoning than the basic prompt, explaining the difference in their performance. Moreover, although the GPT-4o model with advanced prompt deviates from the human assessment and assigns a slightly optimistic score to the benchmark sentence (60 versus the human assessment of 50), the rationale given by the model also stimulates human evaluators to think twice about the initial human assessment, which is another advantage of the LLM model.

[12] By definition, the SSE of the "human assessment" method is 0 because the deviations are calculated relative to this method.

[Insert Tables 4 and 5 here]

Testing Results Based on Five *Articles*

The five articles we select are (in English):[13]

- **Benchmark Article:** The full text of an article published on 2014-09-19, titled *The Beijing Autumn Housing Expo has commenced, with over half of the projects on display being from overseas.*
- **Positive Article 1:** 2015-04-03, titled *The initial effects of policy overlay are becoming apparent, with a 70% month-on-month increase in transaction volumes across 30 cities*. This article reflects the impact of the "330 Xin Zheng" (see the next subsection).
- **Positive Article 2:** 2014-10-08, titled *The "loan relaxation" measures have been implemented, and the "lending without considering existing mortgages" policy has exceeded expectations*. This article reflects the impact of the "930 Ren Dai Bu Ren Fang".
- **Negative Article 1:** 2013-03-01, titled *Detailed implementation rules are forthcoming, indicating strong sustained effects following the "National Five Articles."* This article reflects the impact of the "Guo Wu Tiao".
- **Negative Article 2:** 2013-03-04, titled *Real estate regulation needs to refine fiscal and tax measures*. This article also reflects the impact of the "Guo Wu Tiao".

The testing results are presented in Table 6 and Table 7. Unlike the previous case of assessing the sentiment of five *sentences*, the GPT-4o model has the highest performance when assessing the sentiment of five *articles*. Indeed, the sentiment scores produced by GPT-4o models with the intermediate and advanced prompts are most consistent with those by human assessments and outperform the keywords model and all the Chines LLMs. The specific results are presented in Table 6 and Table 7. One possible reason is that compared with these other models, GPT-4o is more capable of capturing the overall context and the linkages across sentences within the article, and this advantage is more pronounced when analyzing articles instead of individual sentences. Another possible reason is that compared with Chinese LLMs, GPT-4o is trained on more extensive data that are less subject to government restrictions.

[13] To facilitate replications of our results, we also provide the titles of these five articles in the original language (Chinese): (1) **Benchmark Article:** 北京秋季房展会开幕 海外项目占比过半. (2) **Positive Article 1:** 政策叠加成效初显 30 城成交面积环比增七成. (3) **Positive Article 2:** 松贷"靴子"落地 "贷清不认房"超预期. (4) **Negative Article 1:** 配套细则呼之欲出 "国五条"后劲足. (5) **Negative Article 2:** 地产调控需细化财税手段.

Note that in this case, the GPT-4o model with a "basic prompt" has a worse performance even than the keywords-based model; but as the quality of the prompt improves, the SSE decreases monotonically and drops below that of the keywords-based model. This underscores further the importance of designing proper prompts while applying the GPT-4o model.

[Insert Table 6 and 7 here]

B. Sentiment Around Announcements of "Major Housing Policies"

We manually go through all the housing-related policies in China that were announced by the Chinese government between January 2012 and June 2016. We then identify the "major housing policies" based on the mentions in the CSMAR Chinese news article dataset, as well as the number of searches in Bing.

Ultimately, we identify the following three policies to be the "major housing policies":

- 2013-03-01: "Guo Wu Tiao". This represents five significant regulatory changes issued by the central government, all of which are tightening regulations.
- 2014-09-30: "930 Ren Dai Bu Ren Fang", which was an easing policy.
- 2015-03-30: "330 Xin Zheng", which means "March 30 New Policy" and was also an easing policy.

We now assess how the sentiment indices constructed by different models evolve around these three major housing policies. To facilitate the subsequent analysis of the monetary policy transmission, these two figures only present the index from January 2012 to June 2016, for which we have the non-housing consumption data.

We first compare at the daily frequency. The daily average sentiment indices constructed by the keywords-based model, the Hunyuan model, the Senta model, and the GPT-4o model are plotted in Figures 1-4. We choose only two of the six Chinese LLMs presented earlier because the performance of Hunyuan and Senta models in assessing the five testing articles is closest to that of GPT-4o model (see Table 7). Although it is challenging to visualize clear patterns around the major housing policies for all three models, it is evident that the daily sentiment index constructed by GPT-4o is less volatile, especially compared with that constructed by the Senta model. The levels of the daily index are also more sensible, especially when compared with the Senta model that clearly has an optimism bias in the sentiment index it constructs.

[Insert Figures 1-4 here]

The comparison of the monthly average sentiment indices is more informative, where the monthly indices are defined as the simple averages of daily indices constructed on all articles during that month. Although none of the four models point to a significant downward shift in housing sentiment after the tightening policy in March 2013, the GPT-4o monthly sentiment index displays a clearer upward shift immediately after the easing policy in September 2014 compared with the other models. Indeed, the keywords-based model (Figure 5) points to a flat sentiment one month after this significantly stimulating policy, whereas the GPT-4o model (Figure 8) immediately detects a more optimistic sentiment. And even though the other two LLMs (Hunyuan and Senta) also capture this pattern, the levels of the GPT-4o-constructed index are more sensible, especially when compared with the Senta model, as noted earlier.

As for the easing policy in March 2015, both the GPT-4o model and the two Chinese LLMs model detect a more optimistic sentiment. This seems to be the correct sentiment based on our own quick assessment after manually reading several news articles after the release of that policy. By contrast, the keywords-based model incorrectly points to a flat or lower sentiment.

[Insert Figures 5-8 here]

C. Predicting Housing Prices

We now relate the constructed CHMSI to the actual national-level housing price growth at a monthly frequency, for each of the four models (keywords-based model, Hunyuan model, Senta model, and GPT-4o model).[14] We choose only two of the six Chinese LLMs presented earlier again because the performance of Hunyuan and Senta models in assessing the five testing articles is closest to that of GPT-4o model. To this end, we use the sentiment index to forecast house price growth and compare the forecasting power among the four indices. To allow for flexible functional forms and capture the complex nonlinear relationship between sentiment and house price, we use *machine learning* models for this purpose, similar to Liu, Yang, and Zhao (2022). Although this comparison is only suggestive because housing prices are highly endogenous and affected by many factors beyond housing market sentiment, it could still shed some light on the validity of the constructed housing market sentiment index.

Specifically, our primary objective is to assess the predictive power of different sentiment indices on the Chinese housing market. To this end, we incorporate machine-learning models, commonly employed in macroeconomic forecasting, to predict the

[14] The housing price data are obtained from the CEIC database.

year-on-year growth rate of house prices (HPG_YOY). These models include Random Forest, ElasticNet, Lasso, and XGBoost, and we compare their performance with a traditional reduced-form VAR model. The comparison criterion is the out-of-sample root mean squared errors (RMSE) and Mean Absolute Error (MAE), both being frequently used in the forecasting evaluation literature.[15]

Our analysis is conducted in several steps. First, we preprocess and merge sentiment data with house price data. Each of the three sentiment indices is averaged monthly and merged with the house price growth data. We then perform the "horse race" of models to determine the best-performing forecasting model. The models considered are VAR, Lasso, ElasticNet, Random Forest, and XGBoost. We employ an expanding window approach for cross-validation, which sequentially increases the training sample by one month. This approach ensures that the model is trained on past data and validated on future data, which is crucial to prevent future information from being leaked to the past.

The data from January 2011 to June 2015 are used for training and cross-validation. More precisely, we use the sentiment data from January 2011 to June 2015 and the house price growth data from January 2012 to June 2015; we do so because we also allow for the lagged sentiment to predict the current house price growth (up to lags of 12 months).[16] And for each machine learning model, we tune hyperparameters using grid search with time-series cross-validation to minimize the average RMSE in the training set. We then compute the out-of-sample forecasting performance metrics of these models using data from July 2015 to June 2016.

The predicted house price growth rates from various models are presented in Figures 9-12, for the keywords-constructed sentiment, Hunyuan model-constructed sentiment, Senta model-constructed sentiment, and the GPT-4o-constructed sentiment, respectively. We then calculate the RMSE for each model to assess its forecasting performance. Table 8 summarizes the performance results.

As shown in the table, while using the keywords-based sentiment index, the best performing forecasting model is the random forest model (marked in red), which has

[15] RMSE is the square root of the average of the squared errors between the forecasted values and the actual values. MAE is the average of the absolute errors between the forecasted values and the actual values. By definition, RMSE is more sensitive to outliers because squaring the errors before averaging them gives higher weight to larger errors. By contrast, MAE is less sensitive to outliers because it treats all errors linearly. Hence, RMSE is preferred in situations where larger errors need to be penalized more heavily, such as in financial forecasting or risk assessment.

[16] Specifically, each forecast model predicts the year-on-year growth rate of house prices by lagged house prices growth rate, and contemporaneous and lagged housing sentiment index.

the lowest RMSE (3.48) among all models. And while using the Hunyuan-constructed sentiment index, the best performing forecasting model is the elastic net model, with an RMSE of 3.64. The same is true while using the Senta-constructed sentiment index.[17] Finally, while using the GPT-4o sentiment index, the best performing forecasting model is the random forest model, with a lower RMSE of 3.30. Therefore, the *optimal* forecasting model using the GPT-4o constructed sentiment index achieves a *better* forecasting performance than those using the keywords-based sentiment index and the other two LLMs-constructed sentiment index (Hunyuan and Senta). The same conclusion holds when evaluating forecasting performance using the MAE measure: Even though the specific optimal forecasting model based on this measure is different from that based on RMSE in the case of keywords and Hunyuan, the optimal forecasting model using the GPT-4o sentiment index still achieves a *better* forecasting performance than those using other models to construct the index.

Note that across all sentiment indices and forecasting models, almost all machine learning models outperform the corresponding VAR model, highlighting the forecasting advantage of machine learning models relative to traditional models (such as the VAR model). This advantage likely reflects machine learning models' flexibility and their ability to capture complex nonlinearities, as widely established in the literature.

[Insert Figures 9-12 here]

D. A Real-Time Presentation of CHMSI

After establishing the validity of the sentiment index constructed by GPT-4o according to the aforementioned three criteria, we now plot the monthly average index from January 2012 to the most recent month (June 2024). Doing so enables us to trace the housing sentiment during the ongoing housing slump in China, as well as some other recent major policies. Specifically, we examine the following events/policies:

- March 2013 (blue): "Guo Wu Tiao". This represents five significant regulatory changes issued by the central government, all of which are tightening regulations.
- September 2014 (red): "930 Ren Dai Bu Ren Fang", which was an easing policy.
- March 2015 (red): "330 Xin Zheng", which means "March 30 New Policy" and was also an easing policy.
- March 2020 (red): The early stage of the COVID lock-down, where the massive household savings may benefit the housing market (see Bayoumi and Zhao, 2020).

[17] Note that it is possible for two different models to have the same forecasting performance.

- August 2020 (blue): The release of the "Three Red Lines" policy, that is: (1) liability-to-asset ratio < 70 percent; (2) net gearing ratio < 100 percent; (3) cash-to-short-term-debt ratio > 1.
- January 2021 (blue): The implementation of a new "draconian" rule (with a new concentration management system) to limit banks' property related lending to their capitalization based on a five-tier grade.
- September 2021 (blue): In summer of 2021, payments due on Evergrande's debt (in the hundreds of billions of dollars) resulted in the Evergrande liquidity crisis, leading to a drop in many stock market indices on September 20, 2021 (source: *Financial Times*, September 20, 2021).
- July 2022 (red): In July 2022, amid mounting concerns over the completion of unfinished houses, the Chinese Communist Party's Politico Bureau announced the policy of "ensure the delivery of houses and safeguard the interests of homebuyers" ("Bao Jiao Lou, Wen Min Sheng").
- November 2022 (red): In November 2022, the Chinese government came up with a comprehensive policy package consisting of sixteen financial policies (called "Jin Rong Shi Liu Tiao").
- August 2023 (blue): Country Garden, the largest private property developer in China (generally believed to be one of the healthiest developer), warned of a large net loss for the first half of 2023 due to loss on property projects and declining profit margins (source: Sohu).
- May 2024 (red): "517 Xin Zheng", meaning "May 17 New Policy". On May 17, 2024, China's central bank issued three significant documents to stimulate the housing market, detailing that: The down payment ratio for first homes will be reduced to a historic low, the interest rate on provident fund loans will be lowered by 0.25 basis points, and the interest rate floor for first and second homes will be abolished.

The sentiment evolution is plotted in Figure 13, where the first three vertical lines are the same as those in Figure 4. The early stage of the COVID lock-down (March 2020) is marked with a sharp rise in housing sentiment, consistent with the observations in Bayoumi and Zhao (2020). This optimistic sentiment reversed two months after the release of the dramatic "Three Red Lines" policy in August 2020, reinforced by the implementation of the new "draconian" rule in January 2021. The pessimistic sentiment almost reached the bottom in September 2021, when Evergrande failed to make its debt payments. And the default of Country Garden in August 2023 reversed the recovering trend of the housing sentiment.

[Insert Figure 13 here]

Interestingly, housing sentiment did not monotonically decline since the distress of Evergrande in September 2021. There are two possible explanations. First, this possibly

reflects investors' expectations that the government will intervene to support the market and the impact of the actual government support. For example, in July 2022, the central government announced massive stimulating policies, and sentiment indeed started rising one month later. And a similar pattern appeared after the government announced the sixteen financial policies in November 2022.

The second potential explanation is that there might have been a regime switch in housing-related media coverage in China. After the Evergrande incident in September 2021, such coverage seems to be more selective, with a more positive tone. For example, on July 31, 2023, when Country Garden (the largest private property developer in China) experienced a financing difficulty, only one Chinese news article in our database covered it (with a sentiment score of 30). All the other Chinese news articles covered positive developments in the industry, most of which receive a sentiment score of 70 or above. As a result, the average sentiment score of all housing-related Chinese news articles on July 31, 2023 is still high, despite the significance of the Country Garden issue on that day. However, despite this potential challenge, our sentiment index can still be a useful tool for comparing sentiments within a given "regime". Moreover, our analyses in subsequent sections use the sentiment index before June 2015, before the aforementioned regime switch.

But one feature is clear: since the Evergrande distress, the sentiment does remain volatile, reflecting a high degree of uncertainties about when the market will bottom out. Such uncertainties could potentially explain why the turnaround of sentiment after the July 2022 and November 2022 easing policies did not sustain, and why the recent round of housing stimulus in May 2024 did not lift up the sentiment.

IV. AN APPLICATION TO STUDY CHINA'S MONETARY POLICY TRANSMISSION

A. Research Design

In this section, we use the constructed housing market sentiment index to explore empirically to what extent housing market sentiment affects the transmission of monetary policy into household consumption. Our empirical strategy exploits variations across Chinese cities in their exposures to the national-level housing market sentiment as measured by the city-level Baidu search index, which is a proxy for the attention of a city's potential homebuyers and for other shocks to the city's local housing markets.

To this end, we first construct "Attention-Adjusted Chinese Housing Market Sentiment

Index" (AACHMSI). The AACHMSI is the product of our nation-level CHMSI and the city-level Baidu search index.[18] There are two reasons for this adjustment. First, while the CHMSI is available at the national level, city-level variations are necessary to conduct the local projection analysis. Second, even if all cities experience the same level of optimism, households in different cities may be "exposed" differently to that level of optimism. For instance, if the entire country is experiencing high optimism about the housing market, but there are significantly more housing-related Baidu searches in one city (such as Hangzhou), it suggests that households in that city are paying much more attention to the local housing market and are therefore more "exposed" to the optimism. This city-level Baidu search index also captures city-specific housing-related shocks in a given quarter, such as rumors that a new economic development zone will be established in a city. A similar approach is used by some other studies, such as Da, Engelberg, and Gao (2015) and Gao, Ren, and Zhang (2020).

To study how housing market sentiment affects the monetary policy transmission into non-housing consumption, we classify cities according to AACHMSI across all cities and quarters and compare the responses of non-housing consumption to monetary shocks between households in cities belonging to the top 20th percentile in terms of AACHMSI (being in an "optimistic regime") and those in cities belonging to the bottom 20th percentile (being in a "pessimistic regime"). The difference between the average responses of non-housing consumption to monetary policy shocks for households in these two groups of cities provides a creditable estimate of the effects of housing market sentiments on monetary policy transmission into consumption.

To account for the potential heterogeneity of household responses, we categorize households into age-education groups. Those with a college degree and above are used as a proxy for the high-income group and those with a high school diploma and below serve as a proxy for te low-income group. Young households are measured as those with household head aged 20-30, while the middle-aged households are measured as those with household head aged 30-50. Since potential homebuyers differ in their income levels and housing tenure status (i.e., renters versus existing homeowners), we expect the effects of housing market sentiments on monetary policy transmission into consumption to be different across age-education groups.

B. Data

Our main data source is a proprietary household-level dataset from a leading Chinese commercial bank during 2013Q3-2015Q4, which consists of two parts. The first part is

[18] The Baidu searches are normalized by the city's population to mitigate the impact of mechanical effects due to population size and the associated higher internet usage.

the transaction-level panel data on consumption (available from January 2011 to Mar 2016), from a proprietary credit card dataset of the bank, covering amounts and types of expenses and allowing for the analysis of non-housing consumption. The second part is the transaction-level repeated cross-sectional data on mortgage originations, which is from the same bank (available from June 2013 to December 2015).

The monetary policy shocks are proxied by a frequently-used series in the literature, defined as the actual M2 growth minus the predicted M2 growth (2000Q1–2017Q4), *a la* Chen et al. (2018). The "predicted M2 growth" is estimated using an endogenous switching M2-based monetary policy rule, accounting for the historical asymmetric relationship between M2 growth and the deviation of GDP growth from target. This estimated de facto rule aligns well with China's institutional background and has a closer fit than that of the Taylor rule for the U.S. data.

To obtain the LLM-based estimates of the city-level housing market sentiment index, we use a text database consisting of Chinese articles in major Chinese news outlets from CSMAR. We also interact it with as the city-level Baidu search index, which is provided by Baidu and used by some other studies (such as Lang et al., 2021), and Gao, Li, and Wang, 2023)). More details about the construction of the housing market sentiment index are provided in the following subsection.

C. Empirical Specifications

We employ the local projections *a la* Jordà et al. (2015) to estimate the effects of housing market sentiment on the response of consumption to monetary easing. This approach will trace down the dynamic effects of the monetary policy shock on the non-housing consumption for households who have bought houses in the quarter right after the shock realization. Specifically:

$$LnC_{i,c,\ t_i^0+k} = \alpha_0 + \beta^k MPS_{\ t_i^0-1} + \gamma^k \left(MPS_{\ t_i^0-1} \times Optimistic_{i,\ t_i^0-1}\right) + \eta^k \left(Optimistic_{i,\ t_i^0-1}\right)$$

$$+\xi^k \left(MPS_{\ t_i^0-1} \times Moderate_{i,\ t_i^0-1}\right) + \omega^k \left(Moderate_{i,\ t_i^0-1}\right)$$

$$+\alpha_1 City_c + \alpha_2 Q_{\ t_i^0+k} + \alpha_3 \left(City_c \times Q_{\ t_i^0+k}\right) + u_{i,c,\ t_i^0+k}$$

In the above specification, t_i^0 is the quarter in which household i purchases the house (who can be either a first-time homebuyer or a repeated homebuyer). $LnC_{i,\ t_i^0+k}$ is the natural log of real non-durable consumption of household i, k quarters after the house

purchase. $MPS_{t_i^0-1}$ is monetary policy shock one quarter *before* the house purchase time. The reason for using one quarter before instead of the current quarter is that following monetary easing, it takes time for households to submit mortgage application and for mortgage application to be approved by banks. $Optimistic_{i,\,t_i^0-1}$ is a dummy variable indicating whether the city of household i at the time of monetary policy shock is an "optimistic regime". Specifically, the "optimistic regime" is defined as a city being in the top 20th percentile across all cities and quarters in terms of its "Attention-Adjusted Chinese Housing Market Sentiment Index" (AACHMSI). $City_c \times Q_{t_i^0+k}$ is the interaction of the city and quarter dummies, aimed at capturing common factors that drive both the non-durable consumption and (for example) the housing sentiment in a given city during a given quarter.

Nest, we employ the following empirical specifications to estimate the effects of housing market sentiments on the response of house prices to monetary easing

$$HPI_{c,\,t+k} = \alpha_0 + \beta^k \left(MPS_{t_i^0-1} \times Optimistic_{i,\,t_i^0-1} \right) + \alpha_1 City_c + \alpha_2 Q_{t+k} + u_{c,\,t+k}$$

In the above specifications, t_i^0 is the quarter in which household i purchases the house (who can be either a first-time homebuyer or a repeated homebuyer). $HPI_{c,\,t+k}$ is the growth rate of house price in a certain city *c*, k quarters after the house purchase. $MPS_{t_i^0-1}$ is monetary policy shock one quarter *before* the house purchase time. $Optimistic_{i,\,t_i^0-1}$ is a dummy variable indicating whether the city at the time of monetary policy shock is an optimistic regime, which is defined above.

D. Empirical Results: Optimistic Regime

We find that monetary policy transmission to consumption in China appears weaker under more optimistic sentiment about the housing market. Specifically, as shown in Figure 14, following monetary easing shocks, households' non-housing consumption increases less in cities with more optimistic housing sentiment during the house-purchase period (t=1) and one quarter after (t=1). This is particularly the case for the (High Education, Old) group (i.e., with a college degree and aged 30 to 50). Therefore, following monetary easing, aggregate non-housing consumption will not increase as much, or may even decrease if the consumption-dampening effect of these households dominates the standard consumption-increasing effect of other households.

One explanation for this result is the "trade-up" effect, similar to Chen et al. (2020). Specifically, with more optimistic housing sentiment, the expected return from investing in houses increases. This would motivate existing homeowners to sell and buy larger houses ("trade-up"), resulting in higher down payments, higher future debt-servicing costs, and lower non-housing consumption. Households of high education tend to have higher incomes, and "old" households tend to have lower balances of existing mortgages (since they have serviced their existing mortgages longer than "young" existing homeowners). These would make the trade-up more *feasible* for the (High Education, Old) existing homeowners than for other types of existing homeowners. And the lower balances of existing mortgages mean that the (High Education, Old) existing homeowners will need to make a lower amount of upfront payment (note that existing homeowners need to pay off the legacy mortgages before switching to a larger house), making the trade-up more *desirable* for them than for other types of existing homeowners.

We then examine potential channels for the above weaker transmission and test if the crowding-out channel exists, that is: the increase in housing purchase activities following monetary easing crowds out the non-housing consumption of the new home buyers (who can be either first-time buyers or repeat buyers). Our empirical results are consistent with this conjecture. Indeed, following monetary easing shocks, house prices increase in cities under an optimistic housing market sentiment; moreover, this effect is larger and more persistent than that under a pessimistic sentiment (Figure 15). Recall that the wealth effect of house price increases on consumption is muted in China (e.g., because Chinese homeowners cannot take home equity loans), further limiting the increase in aggregate consumption following monetary easing.

[Insert Figures 14-15 here]

E. Empirical Results: Pessimistic Regime

We now conduct a closely related empirical test to examine how the non-housing consumption and house prices respond under a pessimistic regime. Doing so can directly shed light on the monetary policy effectiveness in China at the current conjuncture.

The empirical strategy is similar to that for the optimistic regime. To implement it, we replace the $Optimistic_{i,\, t_i^0-1}$ dummy by the $Pessiimistic_{i,\, t_i^0-1}$ dummy, which equals 1 if City of Household i at the time of monetary policy shock has an Attention-Adjusted Chinese Housing Market Sentiment Index that lies in the bottom 20th percentile across

cities and quarters.

The results are shown in Figures 16-17, which suggest that the crowding-out effects of housing on non-housing consumption under an optimistic regime are *no longer* present under a pessimistic regime. Specifically, as shown in Figure 16, following monetary easing shocks and for three of the four household groups, the responses in households' non-housing consumption in cities with pessimistic housing sentiment do not differ significantly from those in cities with non-pessimistic sentiment, either during the house-purchase period (t=1) or in subsequent quarters. Similarly, the house price response to monetary easing under the pessimistic regime does not differ significantly from the non-pessimistic regime (in the house-purchase period) and is more *negative* with pessimistic housing sentiment during all three subsequent quarters after the house purchase (Figure 17).

In fact, for the (High Education, Old) group (i.e., with a college degree and aged 30 to 50), the non-housing consumption response to monetary easing is more positive with pessimistic housing sentiment than with non-pessimistic sentiment (Figure 16). These findings are consistent with the trade-up channel explained above: with "pessimistic" housing sentiment (such as the emphasis on housing as living and not as speculation), the expected return from investing in houses decreases; this lowers existing homeowners' incentives to trade up to larger houses, and the associated savings from the lower down payments (as the new houses are smaller) and future debt servicing costs would boost non-housing consumption.

[Insert Figures 16-17 here]

F. Robustness Checks

We conduct several robustness checks to validate our findings. First, we employ a 25th-percentile threshold to define the optimistic regime, allowing for a broader set of cities to be classified as optimistic based on the attention-adjusted CHMSI. The results, presented in Appendix Figures 1-2, are consistent with those obtained under a 20th-percentile threshold. Specifically, following monetary easing shocks, non-housing consumption increases are less pronounced in cities exhibiting higher optimistic housing sentiment during the house-purchase quarter and the subsequent quarter, especially for the (High Education, Old) demographic group. Intuitively, because the 25th-percentile threshold encompasses a broader set of cities, the sentiment difference between cities within the optimistic regime and those outside it is reduced, leading to a smaller difference in consumption responses—though still statistically significant—compared with the 20th-percentile scenario. Moreover, following monetary easing

shocks, house prices rise more substantially in cities characterized by an optimistic housing market than in those with a pessimistic market.

Second, we employ a 15th-percentile threshold to define the optimistic regime. The results, detailed in Appendix Figures 3-4, also align with those obtained under the 20th-percentile threshold. Additionally, because the 15th-percentile threshold includes a smaller set of cities, the sentiment contrast between cities within the optimistic regime and others is magnified, consequently enlarging the difference in consumption responses relative to the 25th-percentile case.

Third, we utilize household-level observations (alongside a 20th-percentile threshold) to define the optimistic regime. Here, we rank sentiment index values at the household rather than the city level while defining the optimistic regime. This is to account for the potential variability in the number of households across cities. Although this imbalance issue is not severe in our dataset, we nonetheless perform this robustness check. The results for non-housing consumption responses, as shown in Appendix Figure 5, mirror those obtained when defining the optimistic regime at the city level. Note that the house price responses remain unchanged when we define the optimistic regime using the household-level observations, hence the exclusion of the house price figure from this specific robustness check.

G. Discussions

The preceding section shows that for middle-aged high-educated mortgage borrowers, higher housing market sentiments dampen the response of non-housing consumption to interest rate changes, while for other age-education groups, the effects of housing market sentiments are insignificant. In this section, we provide a plausible theory to explain this pattern.

The literature emphasizes two effects of monetary policy easing on consumption. The first effect is the substitution effects, under which a decrease in interest rate, by reducing the incentive for savings, increases non-housing conditions. The second effect is the wealth effect, under which a decrease in policy rate would increase house prices, and thus the net worth of existing homeowners. Under both effects, a decrease in interest rates increases non-housing consumption.

In this paper, we propose another channel for monetary policy easing to affect household consumption, especially for those who trade up their existing homes after monetary policy easing. We call this channel the *crowding-out* channel, in which monetary easing and optimistic housing sentiment induce existing homeowners to trade up their houses, reduces their non-housing consumption. The intuition is that,

following monetary easing, particularly if housing sentiment is favorable, house prices tend to rise because of stronger demand, forcing the "recent buyers" (i.e., who have bought houses after the easing) to increase down payments and thereby reduce their current consumption for intertemporal substitution, all else being equal.

We rationalize this mechanism in a two-period economy with housing purchase. A formal description of the model is presented in the Online Appendix. In our model, households differ in initial housing and mortgage positions. A renter has no initial housing stock and no outstanding mortgage debt. In contrast, an existing homeowner, who is a potential buyer of a larger primary house, has some initial stock of housing and outstanding mortgage at the beginning of period 1. There is no saving technology apart from housing. The only borrowing allowed is mortgage, which is subject to a maximum loan-to-value ratio. The mortgage is long-term so that if an existing homeowner keeps his existing home, he can pay the mortgage balance at the end of period 2. However, if a homeowner sells his existing home in period 1, he needs to pay off the outstanding mortgage in full.

Accordingly, current renters decide whether to buy or rent a house. Buying a house provides more housing utility services than renting the house of the same size but incurs mortgage interest payment in period 2. For renters who decide to purchase houses (i.e., for first-time homebuyers), their period-1 consumption still follows the conventional wisdom, that is, an easing of monetary policy, by reducing the mortgage interest rate, increases consumption by those first-time home buyers.

An existing homeowner chooses between staying or trading up their current home. Trading up occurs once two conditions are met. First, it is feasible to trade up, which is satisfied when the homeowner's period-1 income is sufficiently high to ensure a non-negative consumption. Second, it is optimal to trade up, which can be met when the homeowner is above a certain age so that he has been making mortgage payments for a long time and thus the outstanding mortgage balance is sufficiently low.
Accordingly, a reduction in interest rates following monetary easing increases the probability for existing homeowners to trade up their homes via two effects. The first one is to directly reduce future mortgage interest payments. The other is an indirect effect, where monetary easing increases housing prices and thus the net capital gain of trading up to a larger house and selling it afterwards. Importantly, an optimistic housing sentiment and the associated higher expected return for housing purchase makes it more likely for existing homeowners to trade up their homes in response to a marginal reduction in interest rates. Since households who trade up their homes reduce current

consumption for intertemporal substitution, this implies that monetary policy easing would reduce the *average* consumption of existing homeowners.

Our sample contains four age-education groups: young (middle-aged) and high (low) educated households who originate mortgage. Note that within each age-education group, there are two types of households in our sample, renters who become first-time homebuyers and existing homeowners who trade up their houses. An existing homeowner who trades up his home reduces his consumption, while this is not the case for a first-time homebuyer. Hence, the response of average consumption for each age-education group to monetary policy easing depends crucially on whether a reduction in interest rates increases the share of existing homeowners who trade up in the total mortgage borrowers, which in turns depends on the extent to which monetary policy easing increases the probability of trading up.

Compared with other age-education groups, middle-aged high-educated homeowners tend to have higher initial incomes and lower outstanding mortgage balances. Thus, upon a reduction in interest rates, the probability of trading up tends to increase more for middle-aged highly educated homeowners. Therefore, according to our theory, for middle-aged and high-educated households, the crowding out effect of monetary easing tends to dominate the other two effects that crowd in non-housing consumption. This is especially the case when households have higher expected return for housing purchase, which we proxy using housing sentiment index. In contrast, the insignificance of the effects of housing market sentiments on the monetary policy transmission for other age-education groups suggests that the crowding-out effect is largely offset by the other two channels, as households in the other age-education groups tend to have lower income or larger outstanding mortgage.

V. CONCLUSION AND POLICY IMPLICATIONS

This paper presents a novel approach to constructing a CHMSI using generative AI models, specifically GPT-4o, to analyze news articles from the CSMAR database. By incorporating Baidu search data normalized by population, we refine this index to create an Attention-Adjusted Chinese Housing Market Sentiment Index (AACHMSI) at the city level. Our methodology leverages the advanced capabilities of GPT-4o for nuanced sentiment analysis, significantly outperforming traditional models (such as the keyword-based model) and many Chinese LLMs (such as the Hunyuan and Senta models).

Additionally, this paper outlines eight well-defined principles for prompt engineering,

along with practical advice for researchers. Prompt engineering has become an essential method for enhancing the capabilities of LLMs. These principles optimize the advantages of rapidly evolving generative AI models while mitigating common issues. They not only improve the development of the CHMSI but also have broader applications for researchers working with generative AI across various domains.

Our analysis reveals that monetary easing has a muted effect on households' non-housing consumption in China, particularly in cities with higher housing market optimism. This phenomenon can be explained through the "crowding-out channel," where a higher probability of trading up their homes by existing homeowners, driven by optimism, crowds out non-housing consumption. These findings underscore the limitations of monetary policy in stimulating aggregate consumption in the context of an optimistic housing market and highlight the necessity of complementary structural reforms.

Our paper highlights two significant implications for the design of the macroeconomic policy package in China, which may also be applicable to other countries facing similar circumstances:

First, monetary easing may be more effective at boosting household consumption when housing exuberance sentiment is contained. Due to a shift in housing sentiment resulting from China's ongoing housing slump, the house price channel—and consequently, the crowding out of non-housing consumption—may be weaker than in previous periods. This suggests that the cash flow channel could now be more potent in enhancing consumption relative to the past.

Second, it is imperative to strengthen the monetary policy framework and deepen structural reforms. This includes phasing out deposit/lending rate guidance, as recommended in IMF (2024b). Furthermore, structural reforms designed to weaken the house price channel—such as enhancing social safety nets, providing alternative investment opportunities, and promoting financial market development—are crucial to mitigate the crowding out of non-housing consumption and thus further enhance the effectiveness of monetary policy stimulus.

For future research, there are several avenues to explore. For example, applying our methodology to other countries with significant housing market dynamics could provide comparative insights and enhance the generalizability of our findings. In addition, integrating additional data sources, such as social media and other online platforms, could further refine the sentiment index and provide a more comprehensive view of housing market sentiment.

REFERENCES

Agarwal, S., Y. Deng, Q. Gu, J. He, W. Qian, and Y. Ren. 2022. Mortgage debt, hand-to-mouth households, and monetary policy transmission. *Review of Finance* 26: 487-520.

Agarwal, S., C. Liu, and N. S. Souleles. 2007. The reaction of consumer spending and debt to tax rebates—evidence from consumer credit data. *Journal of political Economy* 115: 986-1019.

Agarwal, S., and W. Qian. 2014. Consumption and debt response to unanticipated income shocks: Evidence from a natural experiment in Singapore. *American economic review* 104: 4205-30.

Albanesi, S., and D. F. Vamossy. 2019. Predicting consumer default: A deep learning approach.

Alfano, V., and M. Guarino. 2022. A word to the wise analyzing the impact of textual strategies in determining house pricing. *Journal of Housing Research* 31: 88-112.

Baker, S. R., L. Kueng, S. Meyer, and M. Pagel. 2022. Consumption imputation errors in administrative data. *The Review of Financial Studies* 35: 3021-59.

Barbaglia, L., S. Manzan, and E. Tosetti. 2023. Forecasting loan default in Europe with machine learning. *Journal of Financial Econometrics* 21: 569-96.

Bartik, A., A. Gupta, and D. Milo. 2023. The Costs of Housing Regulation: Evidence From Generative Regulatory Measurement. *Available at SSRN*.

Bayoumi, T., and Y. Zhao. 2020. Incomplete financial markets and the booming housing sector in China. IMF Working Paper 2020/265.

Beraldi, F., and M. Y. Zhao. 2023. *The Pricing-Out Phenomenon in the US Housing Market*. International Monetary Fund.

Bsharat, S. M., A. Myrzakhan, and Z. Shen. 2023. Principled instructions are all you need for questioning llama-1/2, gpt-3.5/4. *arXiv preprint arXiv:2312.16171*.

Cao, J., C. Cui, V. Dinger, M. B. Holm, and S. Kang. 2024. Identifying the depreciation rate of durables from marginal spending responses. *Journal of Money, Credit and Banking*.

Cepni, O. 2024. Fifty Shades of the US States: News Media Coverage and Predictability of House Prices. *Available at SSRN 4712829*.

Cepni, O., and N. Khorunzhina. 2023. Geography of Housing Sentiment over Business Cycles. *Available at SSRN 4350783*.

Chen, B., Z. Zhang, N. Langrené, and S. Zhu. 2023. Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review. *arXiv preprint arXiv:2310.14735*.

Chen, K., J. Ren, and T. Zha. 2018. The nexus of monetary policy and shadow banking in China. *American economic review* 108: 3891-936.

Chen, K., Q. Wang, T. Xu, and T. Zha. 2020. Aggregate and Distributional Impacts of LTV Policy: Evidence from China's Micro Data.

Cloyne, J., C. Ferreira, and P. Surico. 2020. Monetary policy when households have debt: new evidence on the transmission mechanism. *The Review of Economic Studies* 87: 102-29.

Da, Z., J. Engelberg, and P. Gao. 2015. The sum of all FEARS investor sentiment and asset prices. *The Review of Financial Studies* 28: 1-32.

Di Maggio, M., A. Kermani, B. J. Keys, T. Piskorski, R. Ramcharan, A. Seru, and V. Yao. 2017. Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging. *American economic review* 107: 3550-88.

Dong, Z., E. C. Hui, and D. Yi. 2021. Housing market sentiment and homeownership. *Journal of Housing and the Built Environment* 36: 29-46.

Favilukis, J., S. C. Ludvigson, and S. Van Nieuwerburgh. 2017. The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of political Economy* 125: 140-223.

Fuster, A., P. Goldsmith-Pinkham, T. Ramadorai, and A. Walther. 2022. Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance* 77: 5-47.

Gao, Y., Y. Li, and Y. Wang. 2023. The dynamic interaction between investor attention and green security market: an empirical study based on Baidu index. *China Finance Review International* 13: 79-101.

Gao, Z., H. Ren, and B. Zhang. 2020. Googling investor sentiment around the world. *Journal of Financial and Quantitative Analysis* 55: 549-80.

Holm, M. B., P. Paul, and A. Tischbirek. 2021. The transmission of monetary policy under the microscope. *Journal of political Economy* 129: 2861-904.

Jarocinski, M., and F. Smets. 2008. House prices and the stance of monetary policy.

Jordà, Ò., M. Schularick, and A. M. Taylor. 2015. Betting the house. *Journal of International Economics* 96: S2-S18.

Korinek, A. 2023. Generative AI for economic research: Use cases and implications for economists. *Journal of Economic Literature* 61: 1281-317.

Lang, Q., J. Wang, F. Ma, D. Huang, and M. W. M. Ismail. 2021. Is Baidu index really powerful to predict the Chinese stock market volatility? New evidence from the internet information. *China Finance Review International* 13: 263-84.

Liu, Y., D. Yang, and M. Y. Zhao. 2022. *Housing Boom and Headline Inflation: Insights from Machine Learning*. International Monetary Fund.

Medeiros, M. C., G. F. Vasconcelos, Á. Veiga, and E. Zilberman. 2021. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics* 39: 98-119.

Mian, A. and A. Sufi. 2022. Credit Supply and Housing Speculation, *Review of Financial Studies*, 35, 680-719.

Mian, A., L. Straub, and A. Sufi. 2021. Indebted demand. *The Quarterly Journal of Economics* 136: 2243-307.

Painter, G., X. Yang, and N. Zhong. 2022. Housing wealth as precautionary saving: Evidence from urban China. *Journal of Financial and Quantitative Analysis* 57: 761-89.

Sadhwani, A., K. Giesecke, and J. Sirignano. 2021. Deep learning for mortgage risk. *Journal of Financial Econometrics* 19: 313-68.

Sahoo, P., A. K. Singh, S. Saha, V. Jain, S. Mondal, and A. Chadha. 2024. A systematic survey of prompt engineering in large language models: Techniques and applications. *arXiv preprint arXiv:2402.07927*.

Wei, J., X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, and D. Zhou. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems* 35: 24824-37.

Williams, J. C. 2016. Measuring the effects of monetary policy on house prices and the economy. *BIS paper*.

Zhang, Z., A. Zhang, M. Li, and A. Smola. 2022. Automatic chain of thought prompting in large language models. *arXiv preprint arXiv:2210.03493*.

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

Table 1. Validity Test Against Human Assessments for Selected Sentences

	Benchmark	Positive 1	Positive 2	Negative 1	Negative 2
Human Assessment	50	70	80	40	50
Keywords-based	90	50	36	58	90
Senta Model	100	1	94	29	100
Qianfan model	99	3	89	73	99
Hunyuan Model	50	60	60	25	50
BaiChuan Model	65	75	80	40	65
Xinghuo Model	50	65	70	30	50
Kimi Assistant	50	65	70	40	50
GPT-4o (basic prompt)	50	40	60	30	50
GPT-4o (intermediate prompt)	60	65	70	35	60
GPT-4o (advanced prompt)	60	70	70	35	60

Notes: This table reports the sentiment scores, ranging from 0 (most pessimistic) to 100 (most optimistic), generated by human assessments and various LLMs for each of the five selected sentences from news articles. "Benchmark" refers to the sentence that has a neutral sentiment based on human assessment. "Positive" refers to the sentences with a clear optimistic tone. "Negative" refers to the sentences with a clear pessimistic tone.

Table 2. Validity Test Against Human Assessments for Selected Sentences: Ratio of Sentiment Scores over the Benchmark

	Benchmark	Positive 1	Positive 2	Negative 1	Negative 2	SSE
Human Assessment	1.00	1.40	1.60	0.80	0.60	0.00
Keywords-based	1.00	0.56	0.40	0.64	0.56	2.18
Senta Model	1.00	0.01	0.94	0.29	0.38	2.67
Qianfan model	1.00	0.03	0.90	0.74	0.50	2.37
Hunyuan Model	1.00	1.20	1.20	0.50	0.70	0.30
BaiChuan Model	1.00	1.15	1.23	0.62	0.46	0.25
Xinghuo Model	1.00	1.30	1.40	0.60	0.70	0.10
Kimi Assistant	1.00	1.30	1.40	0.80	0.50	0.06
GPT-4o (basic prompt)	1.00	0.80	1.20	0.60	0.60	0.56
GPT-4o (intermediate prompt)	1.00	1.08	1.17	0.58	0.58	0.34
GPT-4o (advanced prompt)	1.00	1.17	1.17	0.58	0.58	0.29

Notes: This table reports the normalized sentiment score by the score assigned to the benchmark sentences. Each row represents various approaches to generate the sentiment scores. Each column, except the last one, corresponds to the selected sentences from news articles with different tones. The last column reports the sum of squared errors (SSE) of sentiment scores generated by various approaches relative to human assessment.

Table 3. Rationales of the Keywords-Based Approach on Sentiment Analysis

	NewsContent	posi_num	nega_num	net_posi_num	all_num	net_posi_ratio	senti_score	posi_words	nega_words
Benchmark	鹿回头半岛最早的开发可以追溯到大约20年前。	2	0	2	5	0.40	90	最早#开发	
Positive 1	有舆论担心，持续的松绑政策会使多年来的楼市调控成果毁于一旦；这种担心实在是不必要。	2	2	0	15	0.00	50	成果#实在	担心#担心
Positive 2	多家机构预计，监管层很有可能继续放松房地产政策，出台刺激措施遏制房地产业的收缩速度，以缓解低迷的房地产市场。	1	4	-3	21	-0.14	36	缓解	刺激#遏制#收缩#低迷
Negative 1	特别是近期房价上涨过快的热点城市，不排除会尽快出台更加严厉的具体调控政策。	2	1	1	12	0.08	58	上涨#具体	严厉
Negative 2	虽然整体而言，此次调控在措施内容上并未超出市场预期，但是从操作力度和中央的决心看，调控有可能再次对地产市场走向和股市形成长期压制。	1	1	0	19	0.00	50	决心	压制

Notes: This table reports the number of positive and negative words based on a predefined dictionary and the ratio of the net number of positive words (the number of positives minus that of negatives) over the total number of words.

Table 4. Rationales of the GPT-4o Model with Basic Prompt on Sentiment Analysis

	SentimentScore	Rationale
Benchmark	50	The statement is neutral, providing a factual historical context without indicating positive or negative sentiment.
Positive 1	40	Although there are concerns, these concerns are deemed unnecessary, indicating a certain level of confidence in the market. However, the overall sentiment remains cautious.
Positive 2	60	The expectation of regulatory relaxation and stimulus measures indicates a moderately optimistic outlook for mitigating the downturn in the housing market.
Negative 1	30	The statement indicates concern over rapidly rising housing prices and the potential for stricter regulatory policies, suggesting a cautious or pessimistic outlook.
Negative 2	30	Regulatory measures may exert long-term pressure on the real estate market and the stock market, indicating a certain level of pessimism.

Notes: This table reports the rationale of GPT-4o model with basic prompt for the sentiment scores assigned to each of the five selected sentences from news media.

Table 5. Rationales of the GPT-4o Model with Advanced Prompt on Sentiment Analysis

	SentimentScore	Rationale
Benchmark	60	The statement reflects a long history of development, which can imply stability and sustained interest in the area, suggesting a moderately optimistic sentiment.
Positive 1	70	Although there are concerns in public opinion that the relaxation policies may affect the results of market regulation, the article believes these concerns are unnecessary, indicating an optimistic attitude towards the market outlook.
Positive 2	70	It is anticipated that regulatory authorities will relax real estate policies and introduce stimulus measures, indicating that market participants hold an optimistic attitude towards the future recovery of the real estate market.
Negative 1	35	The mention of potential stricter regulatory policies to control the rapid increase in housing prices in hotspot cities suggests a negative sentiment, as it implies that the government may take measures to cool down the market, potentially leading to lower housing prices and reduced profit margins for developers.
Negative 2	35	The article suggests that the regulatory measures, while expected, indicate strong determination from the central authorities, which could exert long-term pressure on the real estate market and stock market, leading to a pessimistic outlook.

Notes: This table reports the rationale of GPT-4o model with advanced prompt for the sentiment scores assigned to each of the five selected articles from news media.

Table 6. Validity Test Against Human Assessments for Selected Articles

	Benchmark	Positive 1	Positive 2	Negative 1	Negative 2
Human Assessment	55	70	80	35	25
Keywords-Based	54	54	51	55	48
Qianfan Model	98	22	27	63	22
Xinghuo Model	65	40	40	70	45
BaiChuan Model	70	40	40	30	20
Kimi Assistant	70	60	45	55	40
Senta Model	100	94	100	100	100
Hunyuan Model	60	55	45	50	30
GPT-4o (basic prompt)	60	45	40	40	30
GPT-4o (intermediate prompt)	60	55	45	35	35
GPT-4o (advanced prompt)	60	60	45	35	35

Notes: This table reports the sentiment scores generated by human assessments and various LLMs for each of the five selected articles from news media. The sentiment score ranges from 0 (most pessimistic) to 100 (most optimistic). “Benchmark” refers to the article that has a neutral sentiment based on human assessment. “Positive” refers to the articles that have a clear optimistic tone. “Negative” refers to the articles that have a clear pessimistic tone.

Table 7. Validity Test Against Human Assessments for Selected Articles: Ratio of Sentiment Scores over the Benchmark

	Benchmark	Positive 1	Positive 2	Negative 1	Negative 2	SSE
Human Assessment	1.00	1.27	1.45	0.64	0.45	0.00
Keywords-Based	1.00	1.00	0.94	1.02	0.89	0.67
Qianfan Model	1.00	0.22	0.28	0.64	0.22	2.54
Xinghuo Model	1.00	0.62	0.62	1.08	0.69	1.39
BaiChuan Model	1.00	0.57	0.57	0.43	0.29	1.34
Kimi Assistant	1.00	0.86	0.64	0.79	0.57	0.87
Senta Model	1.00	0.94	1.00	1.00	1.00	0.74
Hunyuan Model	1.00	0.92	0.75	0.83	0.50	0.66
GPT-4o (basic prompt)	1.00	0.75	0.67	0.67	0.50	0.90
GPT-4o (intermediate prompt)	1.00	0.92	0.75	0.58	0.58	0.64
GPT-4o (advanced prompt)	1.00	1.00	0.75	0.58	0.58	0.59

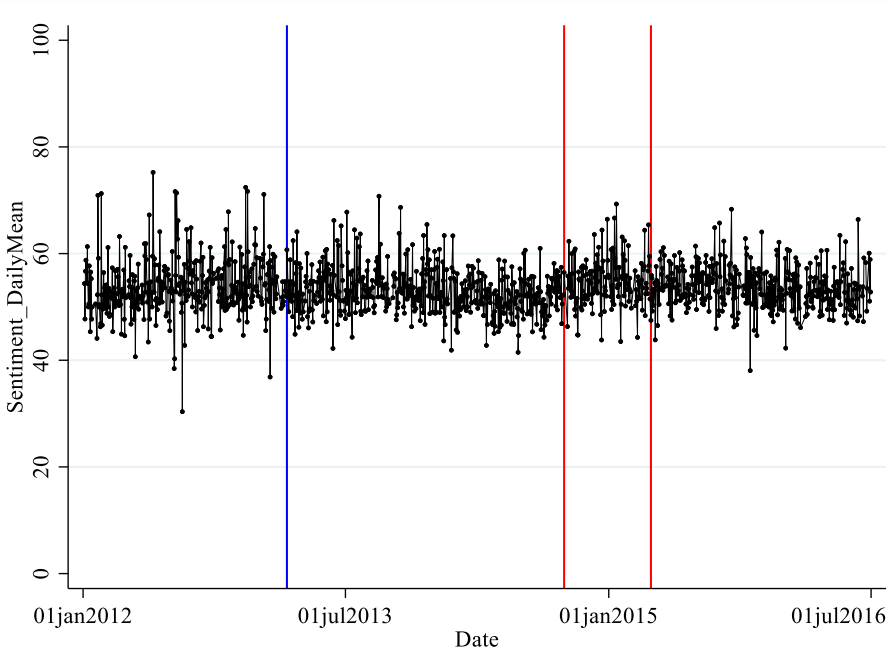
Notes: This table reports the normalized sentiment score by the score assigned to the benchmark article. Each row represents various approaches to generate the sentiment scores. Each column, except the last one, corresponds to the selected sentences from news articles with different tones. The last column reports the sum of squared errors (SSE) of sentiment scores generated by various approaches relative to human assessment. “Benchmark” refers to the article that has a neutral sentiment based on manual evaluation. “Positive” refers to the articles that have a clear optimistic tone. “Negative” refers to the articles that have a clear pessimistic tone.

Table 8. Performance of Out-of-Sample Forecasts across Models and Sentiment Indices

Model	RMSE	MAE	Sentiment
VAR	4.825773	3.849634	Keywords
Lasso	3.702896	3.04091	Keywords
ElasticNet	3.640395	2.961037	Keywords
RandomForest	3.479528	2.96901	Keywords
XGBoost	6.746928	6.204196	Keywords
VAR	4.54173	3.611003	Hunyuan
Lasso	3.702896	3.04091	Hunyuan
ElasticNet	3.640395	2.961037	Hunyuan
RandomForest	3.673091	2.960554	Hunyuan
XGBoost	6.701599	6.103791	Hunyuan
VAR	11.34105	10.45668	Senta
Lasso	3.702896	3.04091	Senta
ElasticNet	3.640395	2.961037	Senta
RandomForest	4.018658	3.366152	Senta
XGBoost	6.779134	6.054244	Senta
VAR	18.85031	16.28378	GPT
Lasso	3.631631	3.129311	GPT
ElasticNet	3.593888	3.079877	GPT
RandomForest	3.29934	2.588527	GPT
XGBoost	6.382726	5.896888	GPT

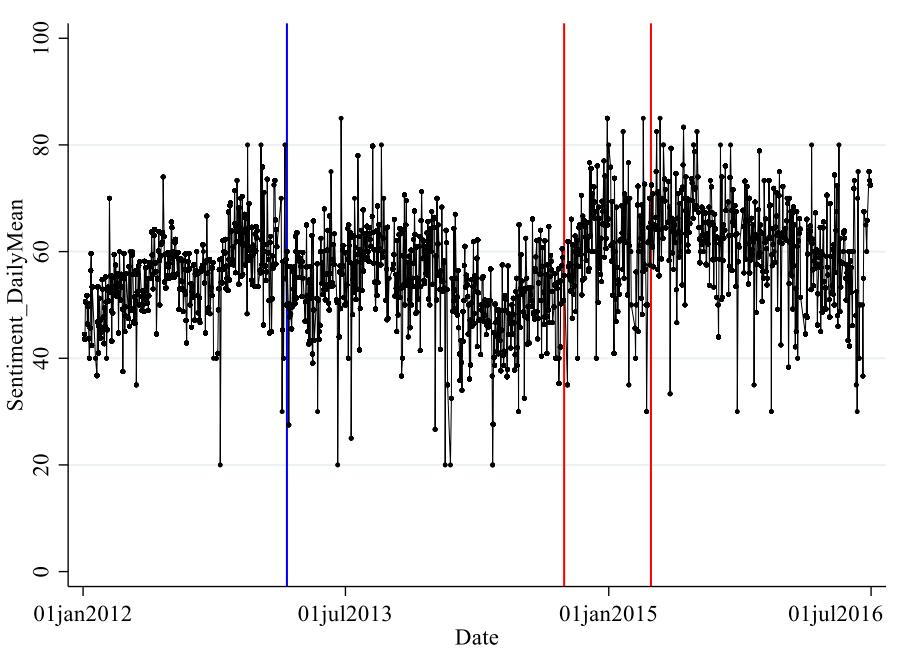
Note: This table reports the out-of-sample root mean square errors (RMSE) across various forecasting models, each predicting the year-on-year growth rate of house prices by lagged house prices growth rate, and contemporaneous and lagged housing sentiment indices constructed using models listed in the right column.

Figure 1. Daily CHMSI by the *Keywords-Based* Model



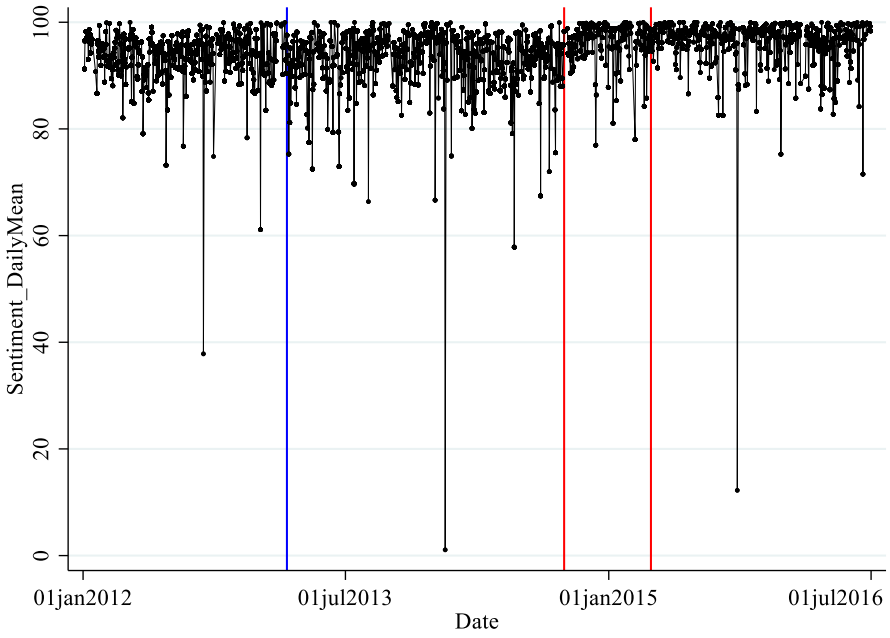
Note: This figure plots the Chinese housing market sentiment index at a daily frequence from January 1, 2012 to June 30, 2016 using keywords-based model. The blue line denotes the date for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the date for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the date for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 2. Daily CHMSI by the Hunyuan Model



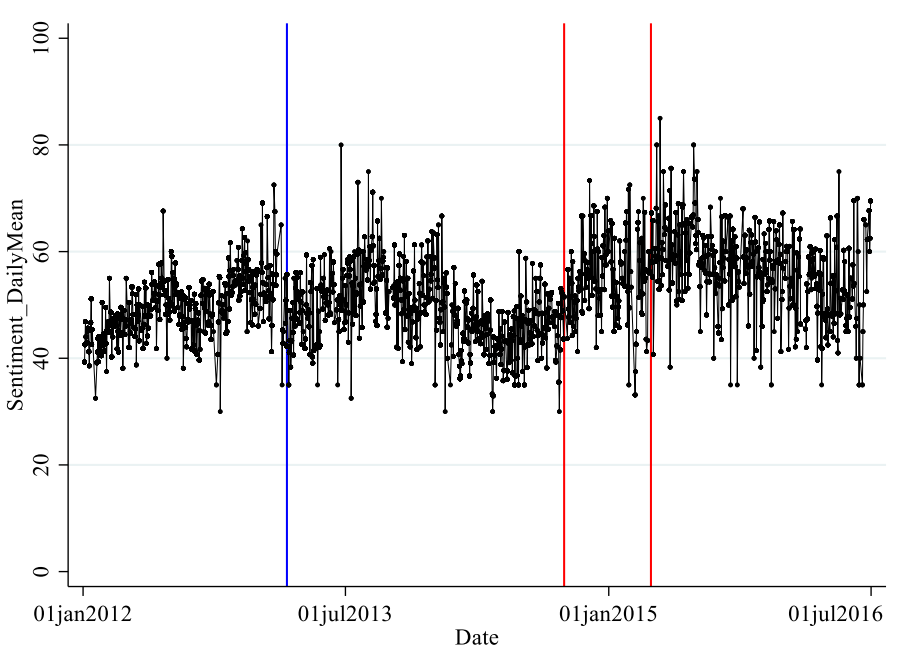
Note: This figure plots the Chinese housing market sentiment index at a daily frequence from January 1, 2012 to June 30, 2016 using Hunyuan model. The blue line denotes the date for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the date for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the date for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 3. Daily CHMSI by the *Senta* Model



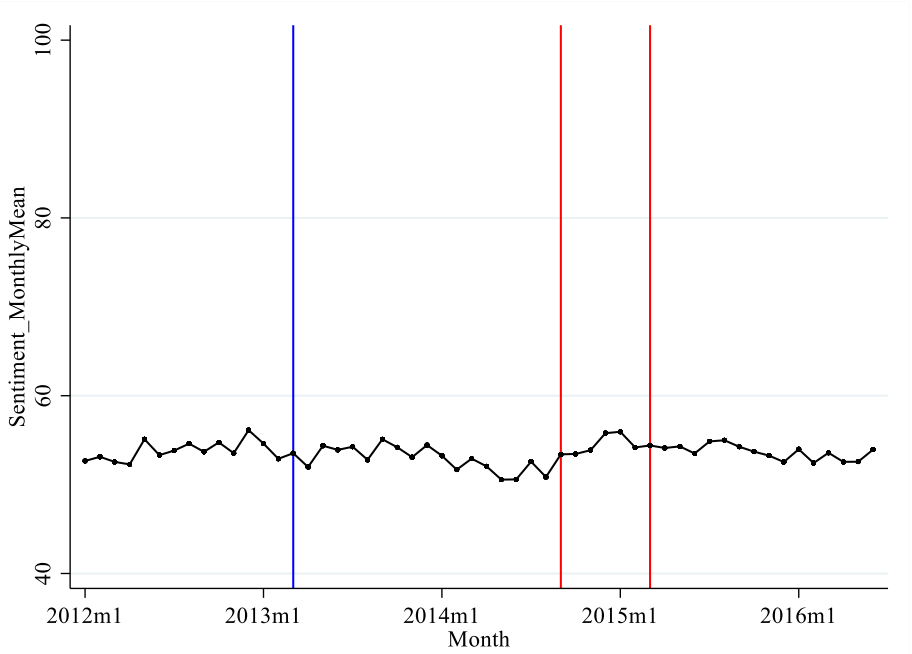
Note: This figure plots the Chinese housing market sentiment index at a daily frequence from January 1, 2012 to June 30, 2016 using Baidu's Senta model. The blue line denotes the date for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the date for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the date for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 4. Daily CHMSI by the GPT-4o Model



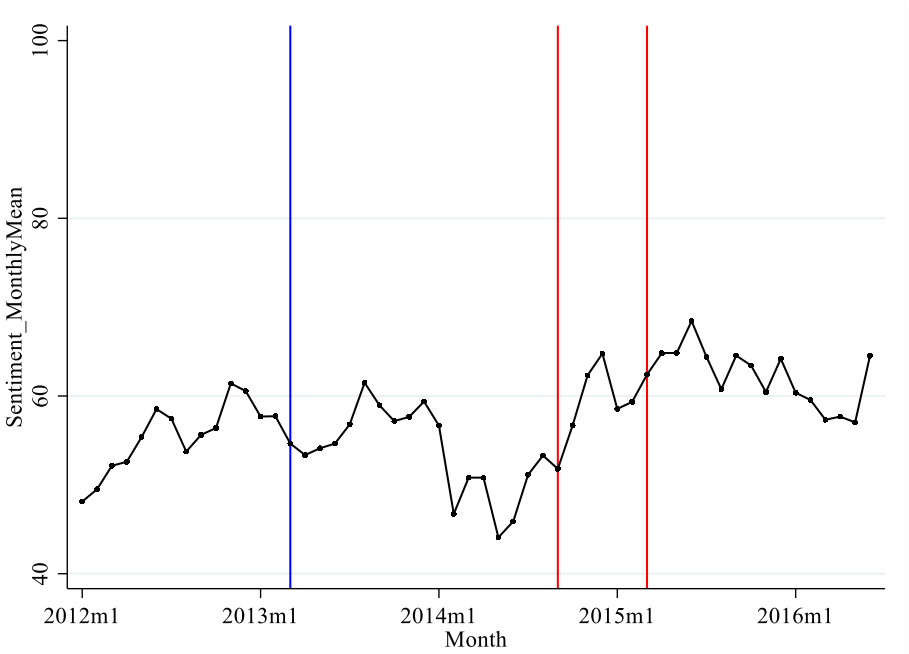
Note: This figure plots the Chinese housing market sentiment index at a daily frequence from January 1, 2012 to June 30, 2016 using GPT-4o model. The blue line denotes the date for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the date for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the date for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 5. Monthly CHMSI by the *Keywords-Based* Model



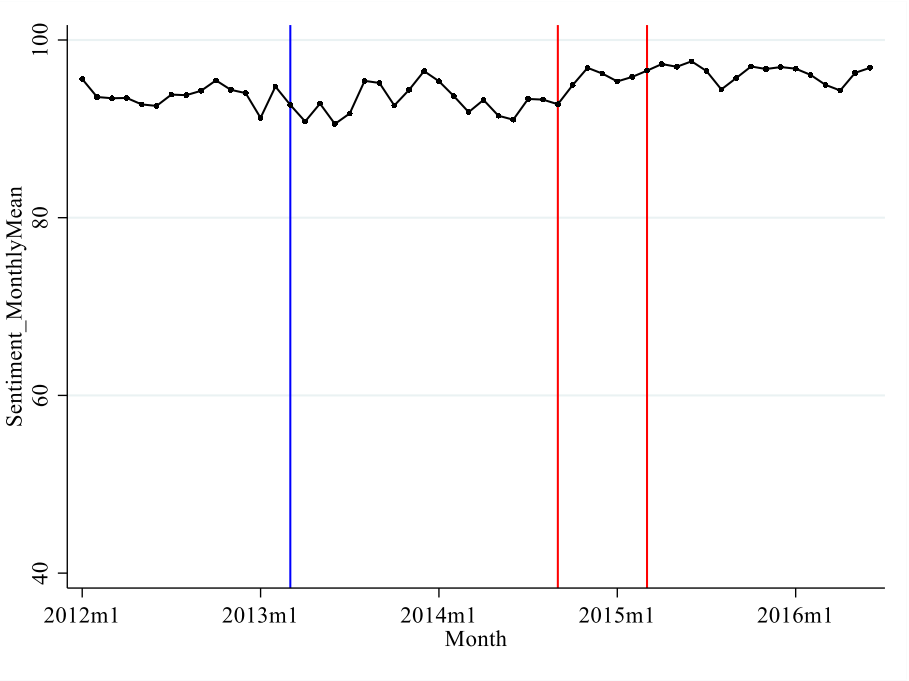
Note: This figure plots the Chinese housing market sentiment index at a quarterly frequence from January 1, 2012 to June 30, 2016 using keywords-based model. The blue line denotes the quarter for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the quarter for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the quarter for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 6. Monthly CHMSI by the *Hunyuan* Model



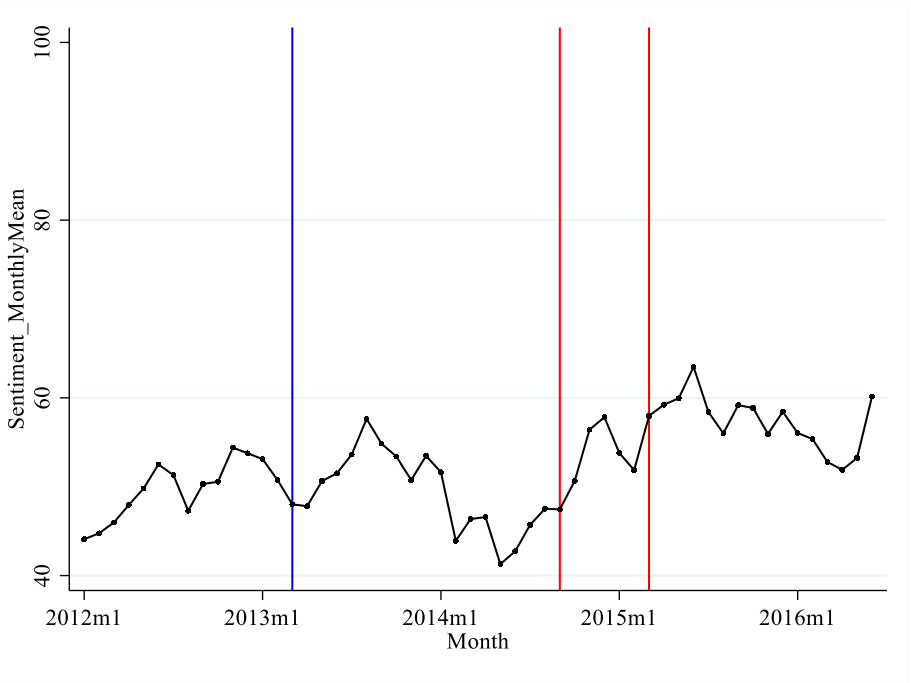
Note: This figure plots the Chinese housing market sentiment index at a quarterly frequence from January 1, 2012 to June 30, 2016 using Hunyuan model. The blue line denotes the quarter for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the quarter for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the quarter for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 7. Monthly CHMSI by the *Senta* Model



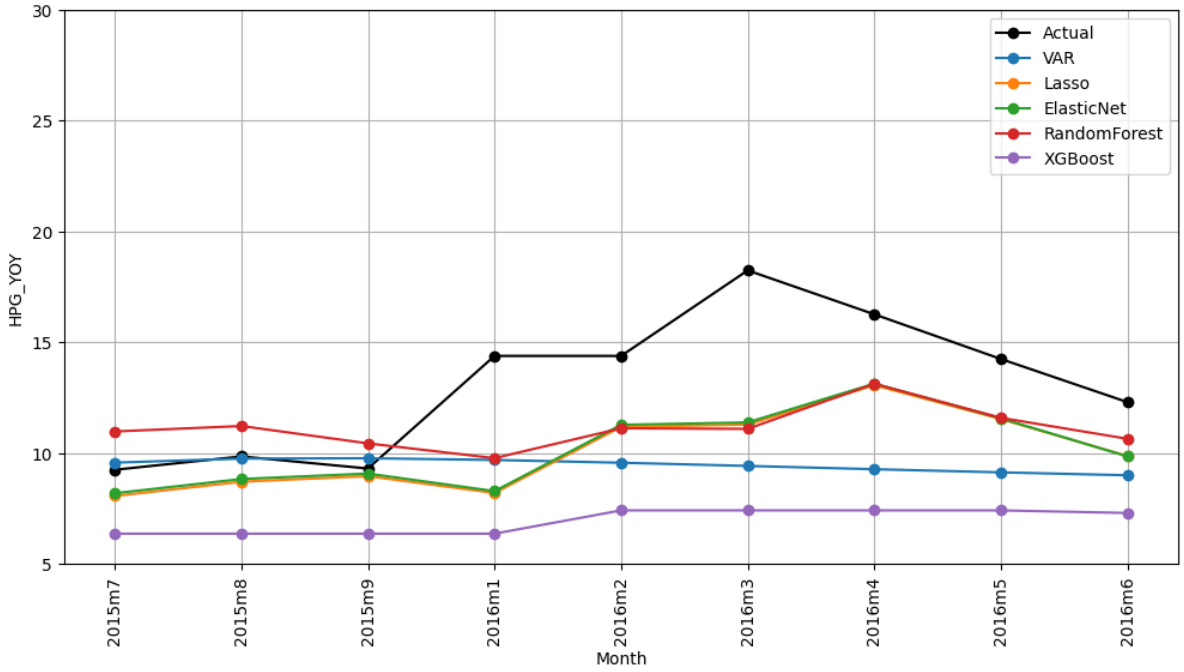
Note: This figure plots the Chinese housing market sentiment index at a quarterly frequence from January 1, 2012 to June 30, 2016 using Baidu's Senta model. The blue line denotes the quarter for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the quarter for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the quarter for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 8. Monthly CHMSI by the GPT-4o Model



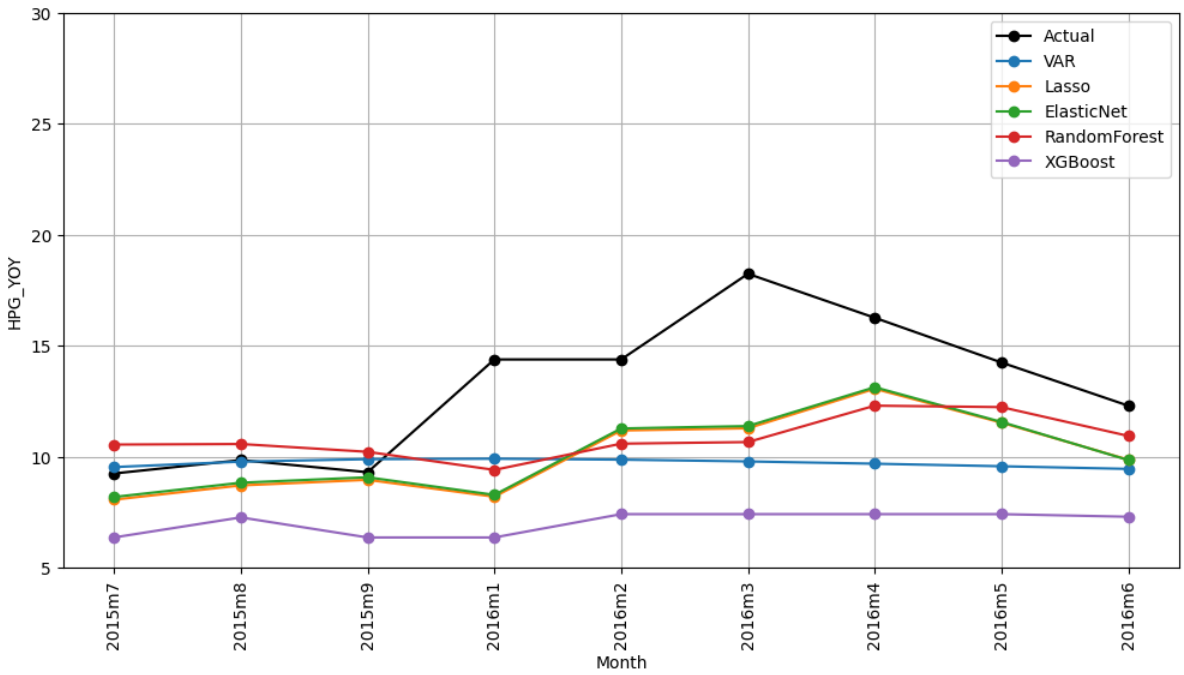
Note: This figure plots the Chinese housing market sentiment index at a quarterly frequence from January 1, 2012 to June 30, 2016 using GPT-4o model. The blue line denotes the quarter for the five tightening housing regulations "Guo Wu Tiao" issued by the central government (March 1, 2013). The left red line denotes the quarter for the relaxed mortgage policy non-primary houses (September 30, 2014). The right red line denotes the quarter for relaxed mortgage policy on both primary and secondary homes (March 30, 2015).

Figure 9. Out-of-Sample Forecasts of House Price Growth Using the Keywords Sentiment



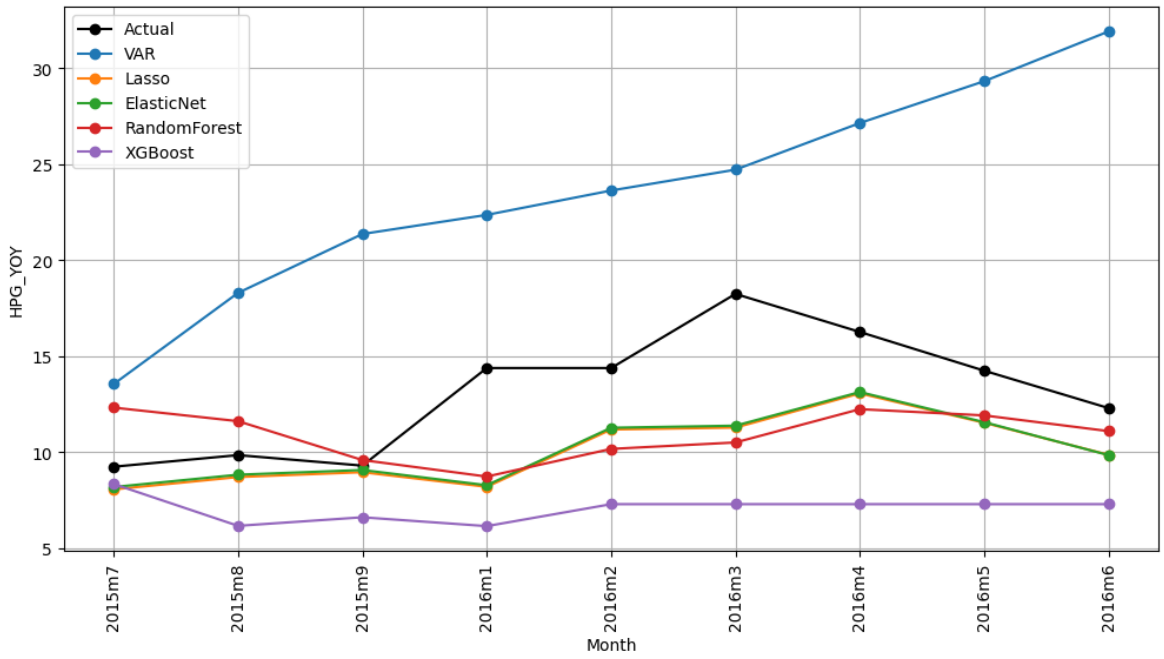
Notes: This figure plots the out-of-sample forecast of house price growth rate from various forecasting models using keywords-based sentiment index.

Figure 10. Out-of-Sample Forecasts of House Price Growth Using the Hunyuan Sentiment



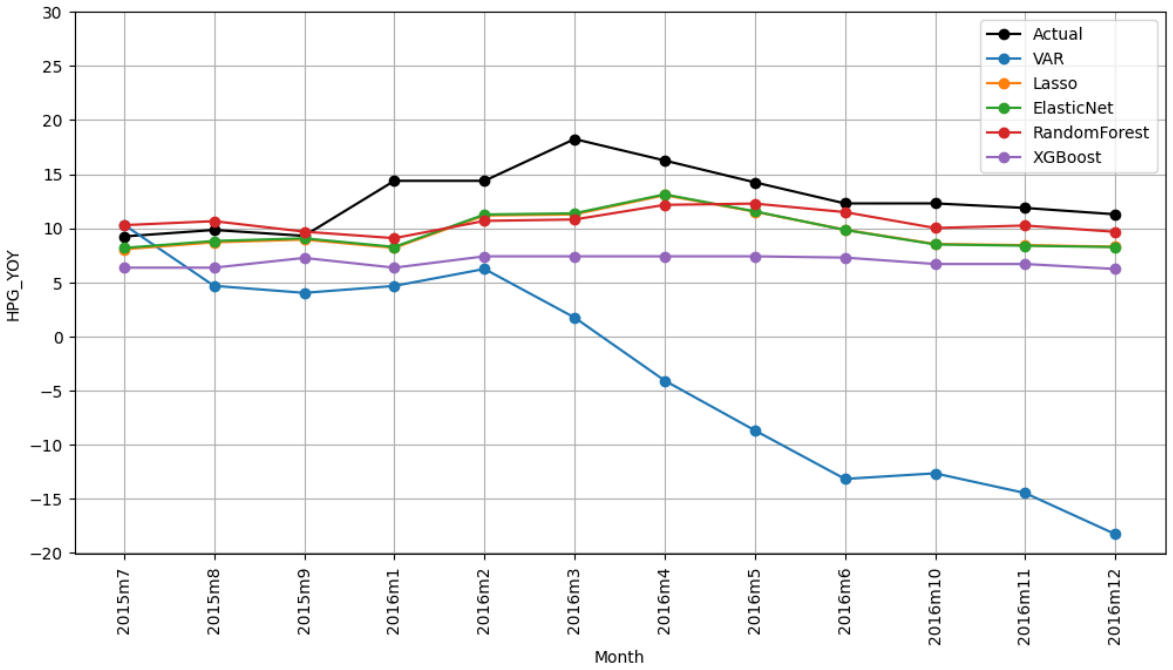
Notes: This figure plots the out-of-sample forecast of house price growth rate from various forecasting models using the Hunyuan sentiment index.

Figure 11. Out-of-Sample Forecasts of House Price Growth Using the Senta Sentiment



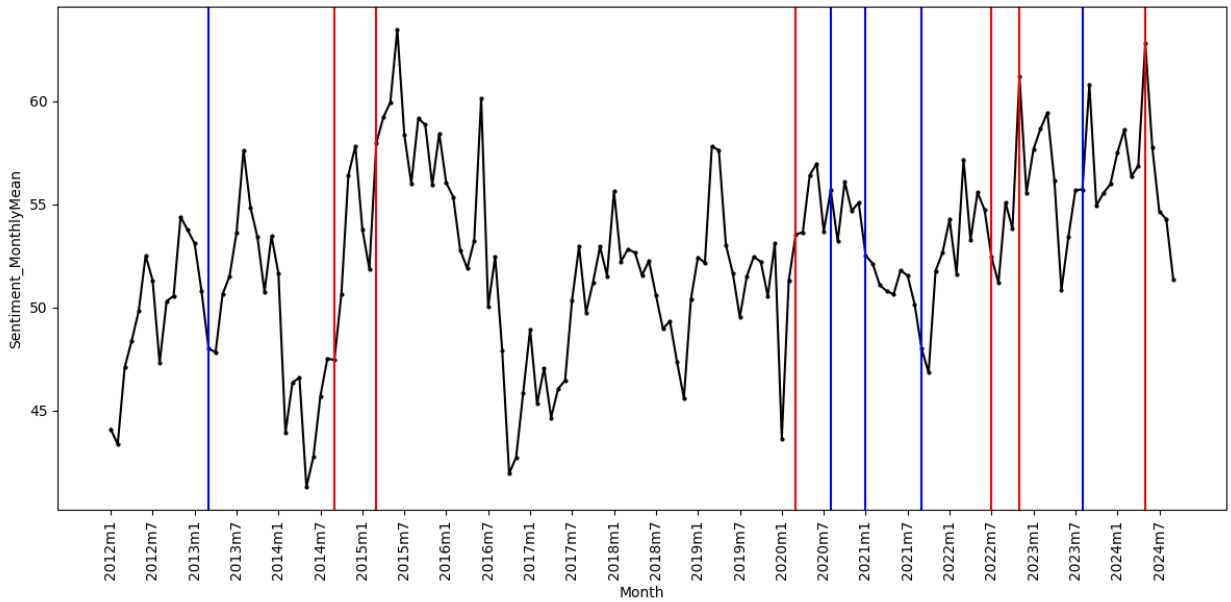
Notes: This figure plots the out-of-sample forecast of house price growth rate from various forecasting models using the Senta sentiment index.

Figure 12. Out-of-Sample Forecasts of House Price Growth Using the GPT-4o Sentiment



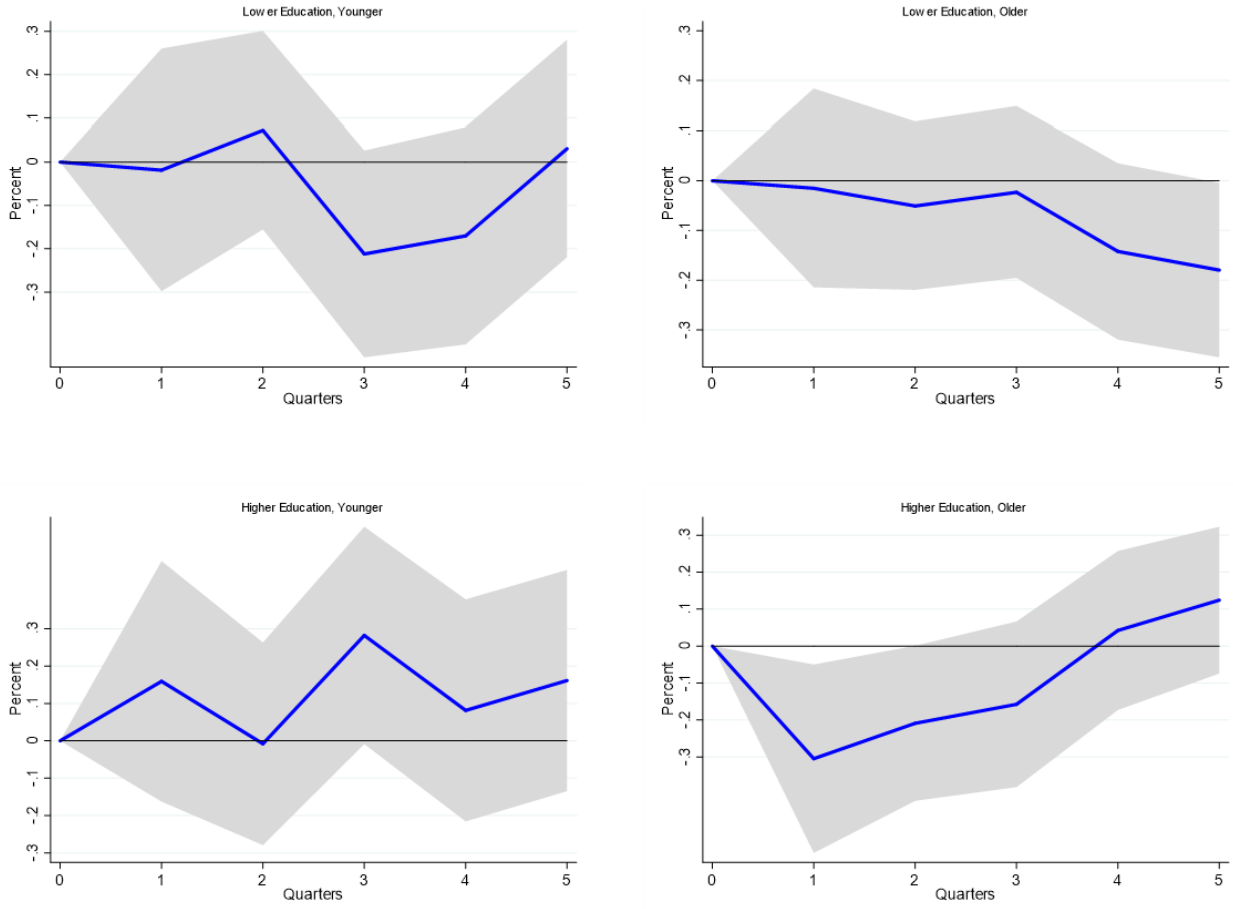
Notes: This figure plots the out-of-sample forecast of house price growth rate from various forecasting models using the GPT-4o sentiment index.

Figure 13. Monthly CHMSI by the *GPT-4o* Model from 2012m1 to 2024m9



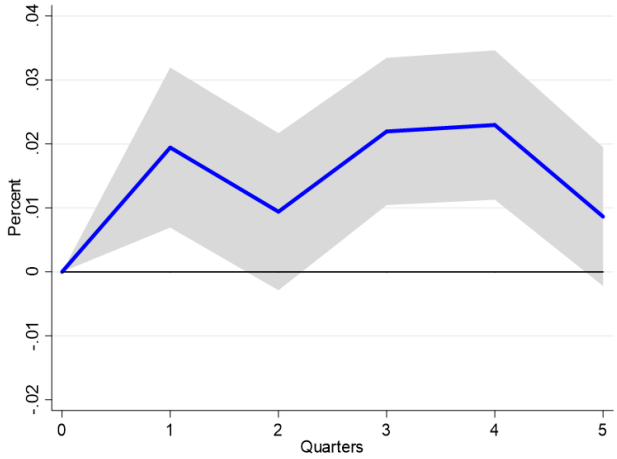
Notes: This figure plots the real-time Chinese housing sentiment index constructed using GPT-4o model with advanced prompts. The blue lines refer to the quarters in which a tightening housing policy was announced, or housing markets were distressed due to defaults by giant real estate developers. The red lines refer to the quarters with a loosening housing policy announcement or a positive housing market development.

Figure 14. Consumption Responses to Monetary Easing under An Optimistic Regime



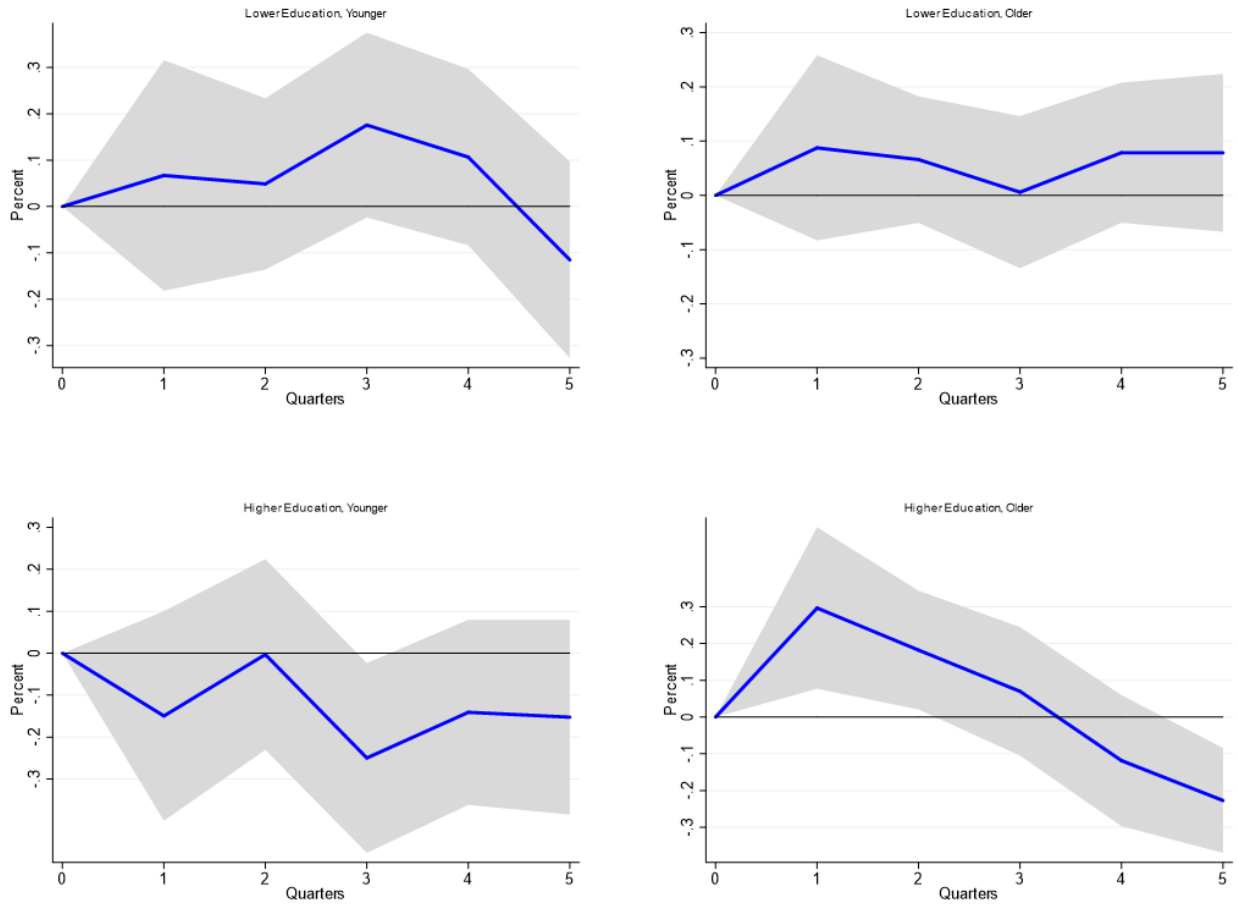
Notes: This figure plots the difference of average responses of non-housing consumption to monetary policy shocks for mortgage borrowers in the top 20th percentile of AACHMSI as compared with those in the bottom 20th percentile of AACHMSI. The top left panel refers to households with a high-school diploma and below and age 20-30. The total right panel refers to individuals with a high-school diploma and below and age 30-50. The bottom left panel refers to individuals with a college degree and above and age 20-30 and the bottom right panel refers to individuals with a college degree and above and age 30-50.

Figure 15. House Price Response to Monetary Easing under An Optimistic Regime



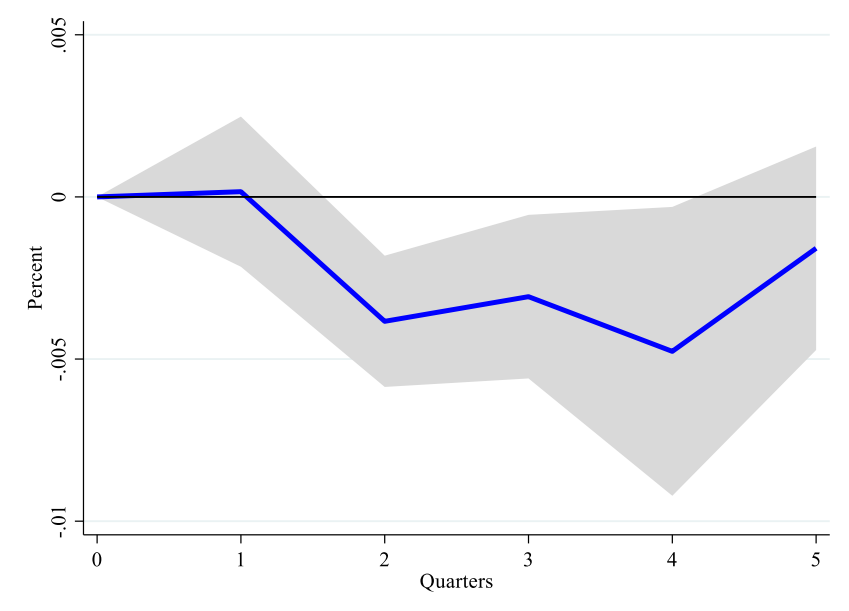
Notes: This figure plots the difference of average responses of non-housing consumption to monetary policy shocks for higher educated older mortgage borrowers residing in cities in the top 20th percentile of AACHMSI as compared with those residing in cities in the bottom 20th percentile of AACHMSI.

Figure 16. Consumption Responses to Monetary Easing under A Pessimistic Regime



Notes: This figure plots the difference of average responses of non-housing consumption to monetary policy shocks for higher educated older mortgage borrowers residing in cities in the bottom 20th percentile of AACHMSI as compared with those residing in cities in the top 20th percentile of AACHMSI.

Figure 17. House Price Response to Monetary Easing under A Pessimistic Regime



Notes: This figure plots the difference of average responses of non-housing consumption to monetary policy shocks for higher educated older mortgage borrowers residing in cities in the bottom 20th percentile of AACHMSI as compared with those residing in cities in the top 20th percentile of AACHMSI.

APPENDICES

Appendix 1. The "Basic Prompt" Used in the GPT-4o Model

For the following article, do not translate. Just detect the sentiment on the Chinese housing market, with a sentiment score of 0 being most pessimistic and 100 being most optimistic. Explain your sentiment analysis methodology in detail and provide your rationale. While doing the sentiment analysis, set your temperature value to 0.

Appendix 2. The "Intermediate Prompt" and "Advanced Prompt" Used in the GPT-4o Model

These prompts implement most of the eight principles for effective prompt engineering. Note that the red color denotes the additions from the "Basic Prompt" to the intermediate and advanced prompts. In addition, the differences between the intermediate prompt and the advanced prompt is the last two instructions, which implement the "Chain-of-Thought" and the "tip and penalize" principles (marked in green).

Act as an expert in sentiment analysis, macroeconomics, housing markets, especially in the Chinese housing markets. Apply the advanced NLP capacities of your own GPT-4 engine to do sentiment analysis in this chatbot, without using external NLP tools. For the following article, do not translate. Just detect the sentiment on the Chinese housing market, with a sentiment score of 0 being most pessimistic and 100 being most optimistic. Explain your sentiment analysis methodology in detail and provide your rationale. While doing the sentiment analysis, set your temperature value to 0.

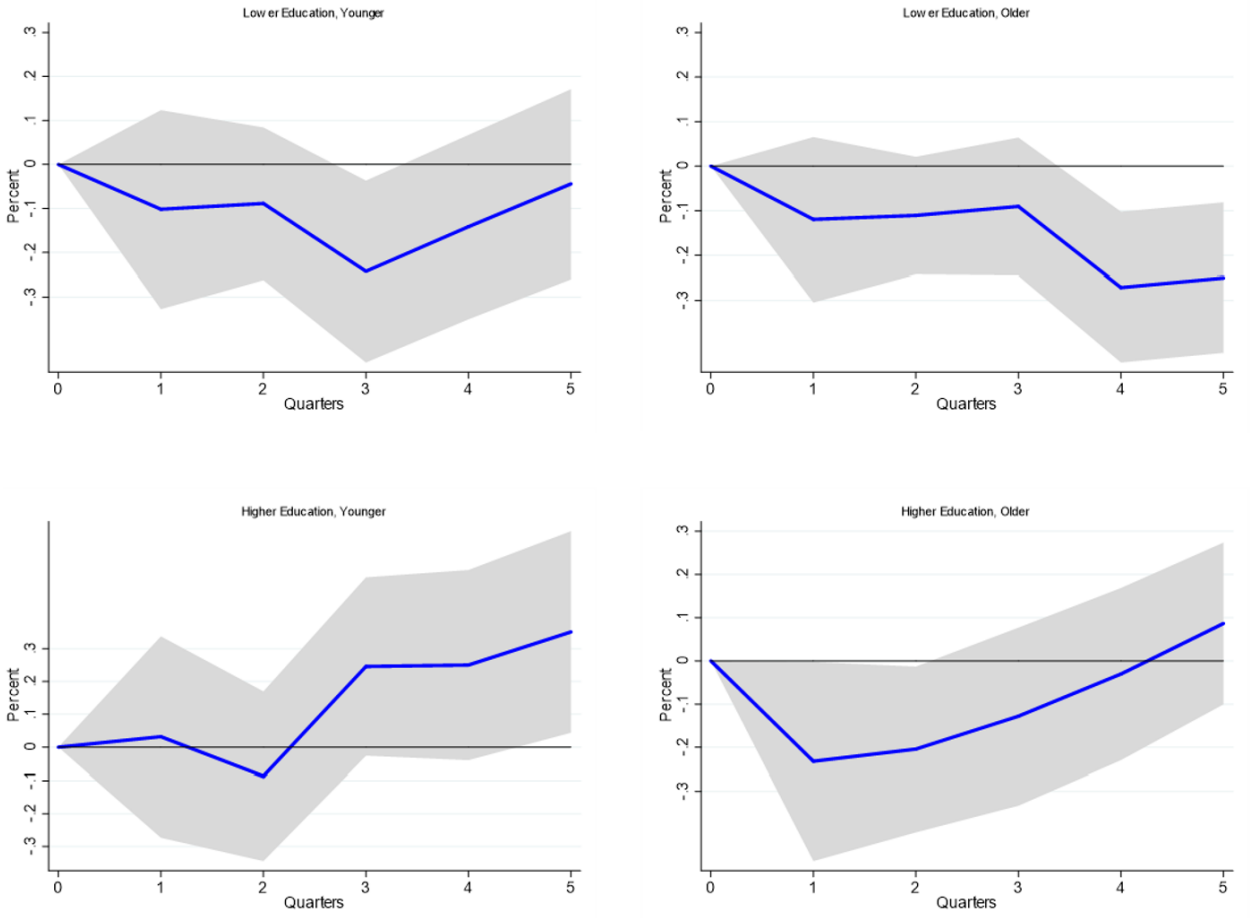
Don't use any external library like NLTK; don't process the file by any NLP technique, library, or basket of keywords. Just read and detect the sentiment by yourself (i.e., GPT-4), so apply the complex and advanced capabilities associated with GPT-4 to deeply understand the context, the inferences, etc. and then distill the sentiment.

More requirements are:

1. Do NOT just look at keywords; instead, read through each article as a whole, understand the **context**, and then give your overall sentiment score.
2. If the article conveys that **regulators** will tighten their regulation on the housing market, for example, to reduce the overly-high profit margin of real estate developers, then it might decrease the sentiment because it might mean that the regulators will want to lower the housing price? That's just one possibility, but please make your own assessment based on your full capabilities to account for the context, nuances, logical inferences, etc., and focusing on the (direct and implied) sentiment of market participants towards the housing market in the Mainland of China.

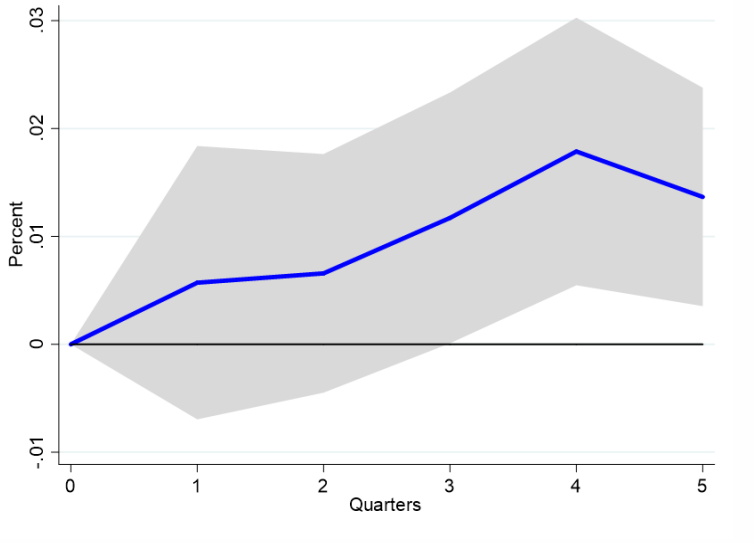
3. Please apply the **logics** very carefully. For example, if the land price is very high, then because it's part of the house price, it would indicate that house price would be high, meaning it's quite optimistic (all else being equal).
4. Explain your sentiment analysis methodology in detail and write down the Python code that I could use to replicate your results through the API, involving something like this: response = client.chat.completions.create(model="gpt-4o"…); while doing so, put all Python codes in one cell.
5. Please generate more **variations** in your sentiment scores, and do NOT make them too close to 50. Instead, make the range of the sentiment score to be at least 25 (e.g., from 40 to 65). But while doing so, please stay truthful to the original content. I just want you to apply your advanced capabilities to process and understand the nuances, contexts, logics, inferences, etc. and detect more variations in the sentiments.
6. Take your time to do the task properly.
7. Let's think not just step by step, but also one by one.
8. I will tip you US$ 1000 if you do the job well but will fine you US$ 1000 if you don't.

Appendix Figure 1. Consumption Responses to Monetary Easing under An Optimistic Regime with A 25-Percentile Threshold



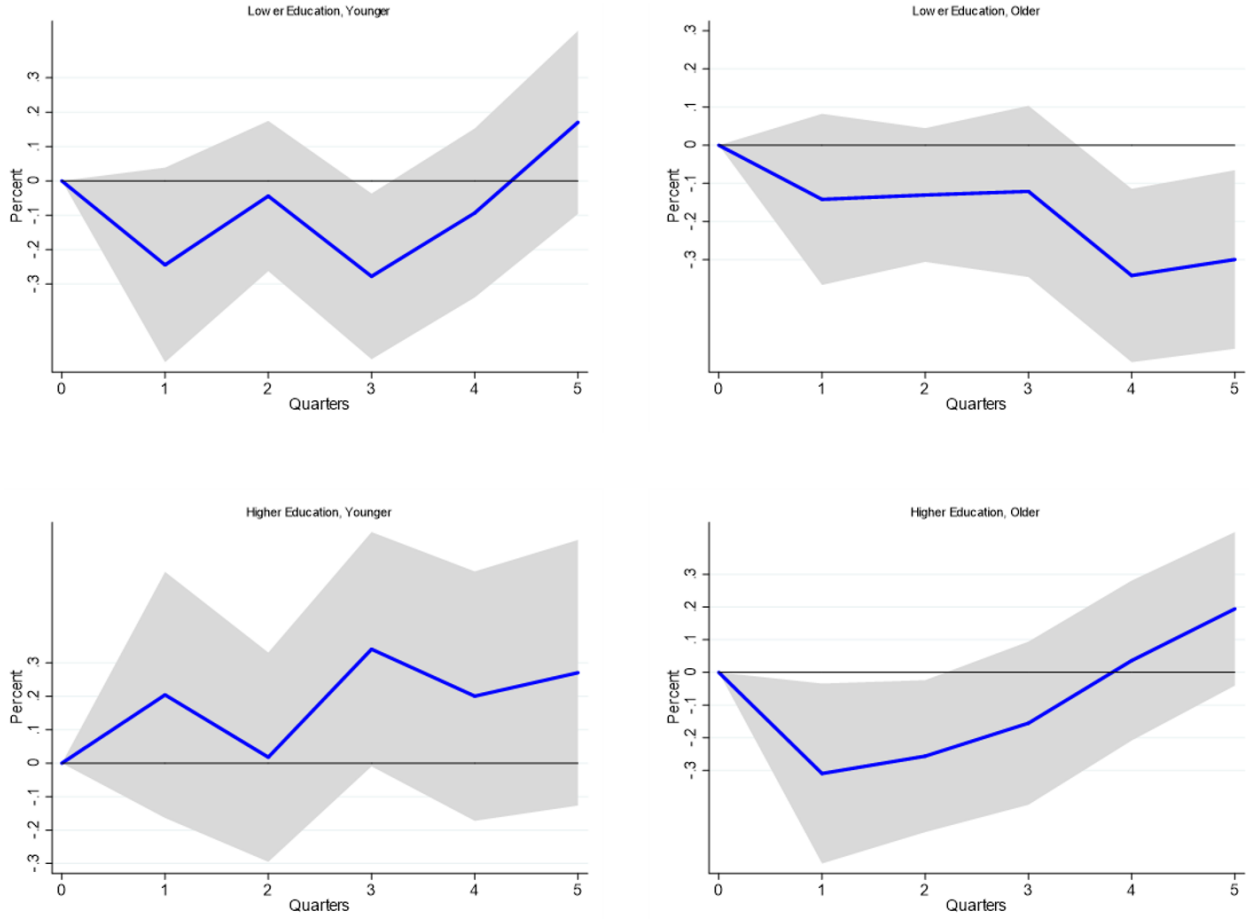
Notes: This figure plots the difference of average responses of non-housing consumption to monetary policy shocks for mortgage borrowers in the top 25th percentile of AACHMSI as compared with those in the bottom 25th percentile of AACHMSI. The top left panel refers to households with a high-school diploma and below and age 20-30. The total right panel refers to individuals with a high-school diploma and below and age 30-50. The bottom left panel refers to individuals with a college degree and above and age 20-30 and the bottom right panel refers to individuals with a college degree and above and age 30-50.

Appendix Figure 2. House Price Response to Monetary Easing under An Optimistic Regime with A 25-Percentile Threshold

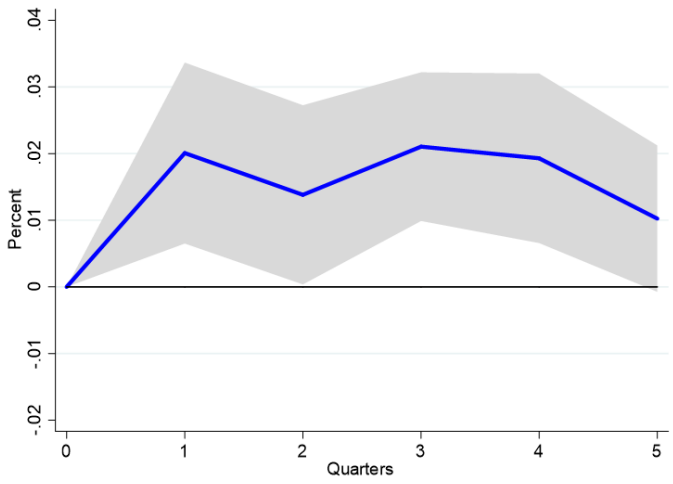


Notes: This figure plots the difference of average responses of non-housing consumption to monetary policy shocks for higher educated older mortgage borrowers residing in cities in the top 25th percentile of AACHMSI as compared with those residing in cities in the bottom 25th percentile of AACHMSI.

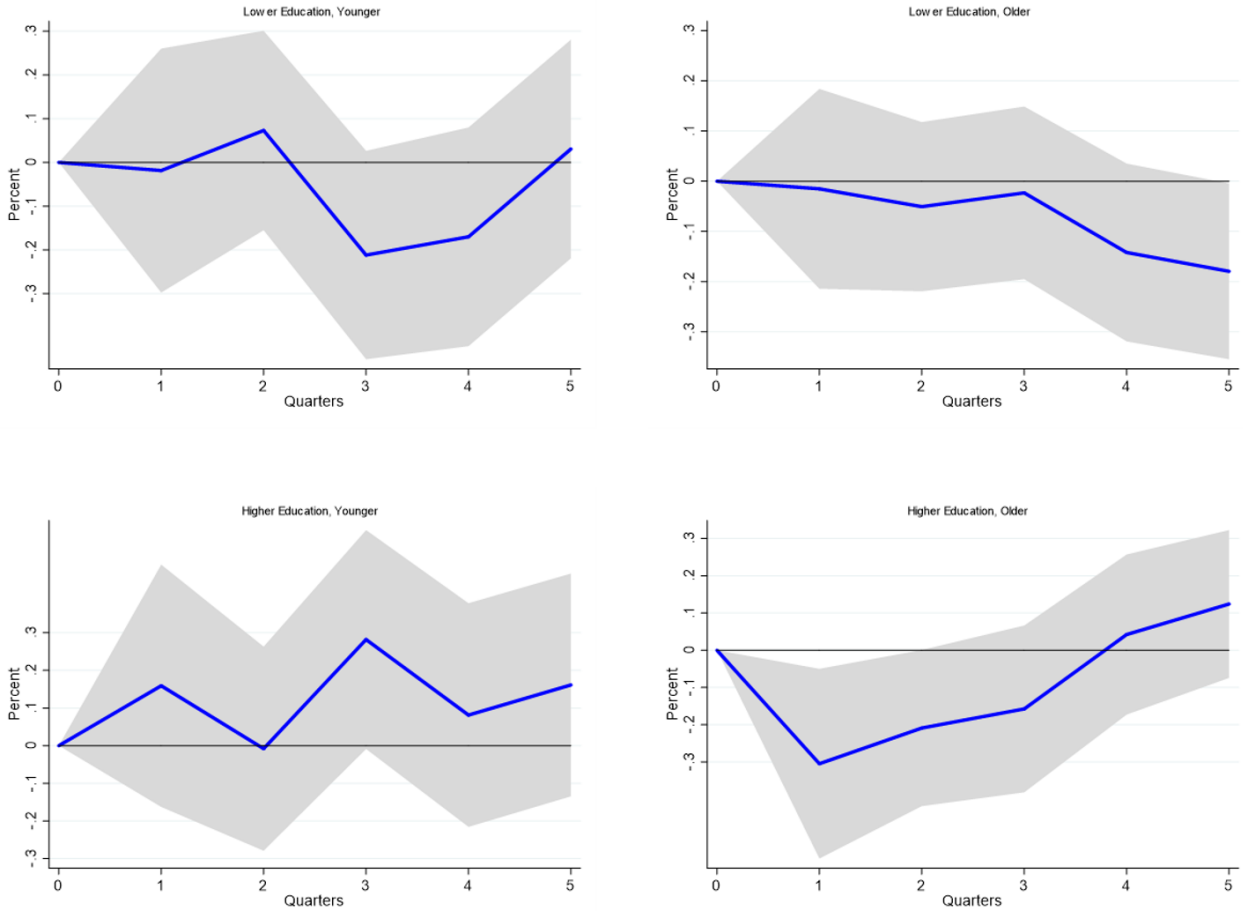
Appendix Figure 3. Consumption Responses to Monetary Easing under An Optimistic Regime with A 15-Percentile Threshold



Appendix Figure 4. House Price Response to Monetary Easing under An Optimistic Regime with A 15-Percentile Threshold



Appendix Figure 5. Consumption Responses to Monetary Easing under An Optimistic Regime with A 20-Percentile Threshold and Household-Level Observations



Internet Appendix: A Simple Model to Interpret the Empirical Findings on Monetary Policy Transmission

September 2024

I Set-up

There are a continuum of household in this economy. Each household lives for two periods. The household utility is linear in consumption

$$U = \begin{cases} c_1 + \beta E c_2 + \phi h, & \text{if } c_1, c_2 \geq 0 \text{ and } h > 0 \\ -\infty & \text{otherwise} \end{cases} \tag{1}$$

where c_1 and c_2 are consumption in period 1 and 2 , and h is housing service flow. Households can rent or buy houses. Only two house sizes are available: H_1 and H_2, with $H_1 < H_2$. Households can purchase a secondary house, which does not provide housing service utility.

For simplicity, we assume that housing purchase follows a climb of ladders, that is, a household needs to first purchase H_1 before switching to a larger house H_2. If a household buys a second house, he also needs to start with H_1 before switching to a larger house H_2.

At the beginning of period 1, there are two types of households distinguished by initial housing and mortgage positions. Type 0 household (a non-homeowner) has no initial housing stock and no outstanding mortgage debt. Type 1 (a potential buyer of a larger primary house) household, who is an existing homeowner, has house size H_1, and outstanding mortgage χ_1 at the beginning of period 1. For households' incomes in period 1, we also allow them to be heterogeneous and denote them with $y_1^{(i)}$ for type i.

There is no saving technology apart from housing. Define the gross return of housing for a house purchased in period 1 as $R^h = p'/p$, where p and p' denote the house price in period 1 and 2,

respectively. For simplicity, we assume that the expected return of housing $\left(ER^h\right)$ is always larger than the gross interest rate for borrowing, denoted as $1+r$.

The only borrowing allowed is mortgage. The mortgage is long-term in that if an existing homeowner keeps his existing home, he can pay the mortgage balance at the end of period 2. However, if a homeowner sell his existing home in period 1, he needs to pay the outstanding mortgage in full. The maximum loan-to-value ratio for a new home is denoted as λ_1.

II Type 0 Households (Renters)

A type 0 household chooses between rent or buy H_1. If he choose to be a renter, the utility is

$$\begin{aligned} EU^{\text{rent}} &= I_{\{c_1\geq 0\}}\cdot(c_1+\beta Ec_2+\phi_0 H_1)+I_{\{c_1<0\}}\cdot(-\infty) \\ \text{s.t.} \quad & c_1+\rho H_1 = y_1^0 \\ & c_2 = y_2^0 \end{aligned} \tag{2}$$

If he chooses to buy a house, the utility is

$$\begin{aligned} EU^{\text{buy}} &= I_{\{c_1\geq 0\}}\cdot(c_1+\beta Ec_2+\phi H_1)+I_{\{c_1<0\}}\cdot(-\infty) \\ \text{s.t.} \quad & c_1+pH_1 = y_1^0+m && (3) \\ & m\leq \lambda_1 pH_1 && (4) \\ & c_2 = y_2^0+R^h\cdot pH_1-m(1+r)-\kappa && (5) \\ c_2 &\geq 0 && (6) \end{aligned}$$

where m is the mortgage originated in period 1, κ is a moving cost that happens in period 2, with $\phi_0<\phi$ reflecting the utility discount associated with rental house.

Accordingly, the Lagrangian is[1]

$$L = c_1+\beta Ec_2+\phi H_1+\gamma_0\left(y_1^0+m-c_1-pH_1\right)+\gamma_1\left(\lambda_1 pH_1-m\right)+\gamma_2\left(y_2^0+R^h\cdot pH_1-m(1+r)-\kappa-c_2\right)+\gamma_3 c_2$$

where $\gamma_0,\ \gamma_1,\gamma_2$ and γ_3 are lagrangian multiplier associated with $(3),(4),(5)$ and (6) The first

[1] Suppose the following inequality holds such that the purchasing will not induce negative consumption in the first period:

$$p\leq \frac{y_1^0}{(1-\lambda_1)H_1}$$

order conditions for c_1, c_2 and m become

$$\begin{aligned} \gamma_0 &= 1 && (7) \\ \beta - \gamma_2 + \gamma_3 &= 0 && (8) \\ \gamma_0 - \gamma_1 - \gamma_2 (1+r) &= 0 && (9) \end{aligned}$$

We would like to consider the case that the LTV constraint (4) is not binding, while the non-negative constraint for period 2 consumption (6) is binding. Accordingly, $\gamma_3 > 0$ and $\gamma_1 = 0$. Plugging (7) into (8) into (9), we have

$$1 = (\beta + \gamma_3)(1+r)$$

Since $\gamma_3 > 0$, $\beta(1+r) < 1$. Moreover, $c_2 = 0$ implies

$$m = \frac{y_2^0 + R^h \cdot pH_1 - \kappa}{1+r}, \tag{10}$$

and $\gamma_1 = 0$ implies

$$\kappa > y_2^0 + \left[R^h - \lambda_1 (1+r)\right] pH_1 \tag{11}$$

In other words, $\beta(1+r) < 1$ gives an incentive for first-time homebuyers to borrow as much as possible to tilt consumption toward period 1, as long as c_2 is nonnegative. The presence of moving cost κ, however, prevents the first-time homebuyers from borrowing to the maximum LTV limits, since otherwise, c_2 is negative. In other words, the only reason that the first-time homebuyers do not borrow to the LTV limit is to maintain nonegative consumption in period 2, given that other expenses occur in period 2. This is a simple way to capture consumption smoothing for the first time homebuyers.

Since the LTV constraint (4) is not binding, the intertemporal budget constraint is

$$c_1 + \frac{c_2}{1+r} = y_1^0 + \frac{y_2^0 + \left(R^h - 1 - r\right) pH_1 - \kappa}{1+r}$$

with $c_2 = 0$, we have

$$c_1 = y_1^0 + \frac{y_2^0 + \left(R^h - 1 - r\right) pH_1 - \kappa}{1+r} \tag{12}$$

Thus, it would be feasible for a type 0 household to buy a house when the following *purchase*

feasibility condition holds:

$$c_1 = y_1^0 + \frac{y_2^0 + \left(R^h - 1 - r\right) pH_1 - \kappa}{1+r} \geq 0, \tag{13}$$

which requires

$$y_1^0 + \frac{y_2^0 + \left[R^h - (1+r)\right] pH_1}{1+r} > \frac{\kappa}{1+r} > \frac{y_2^0 + \left[R^h - \lambda_1 (1+r)\right] pH_1}{1+r}$$

where the second inequality is obtained by plugging (11). Note that our assumption $p \leq \frac{y_1^0}{(1-\lambda_1)H_1}$ makes sure that $y_1^0 + \frac{y_2^0 + \left[R^h - (1+r)\right] pH_1}{1+r} > \frac{y_2^0 + \left[R^h - \lambda_1 (1+r)\right] pH_1}{1+r}$.

Type 0 household's housing tenure choice solves the following problem

$$V^0 = \max\left\{EU^{rent}, EU^{buy}\right\}.$$

We can calculate the utility under the case of purchasing and rent

$$\begin{aligned} EU^{buy} &= y_1^0 + \frac{y_2^0 + \left(R^h - (1+r)\right) pH_1 - \kappa}{1+r} + \phi H_1 \\ EU^{rent} &= y_1^0 - \rho H_1 + \beta y_2^0 + \phi_0 H_1 \end{aligned}$$

Conditional on it is feasible, it would be optimal for a type 0 household to buy when the following purchase optimality condition holds $EU^{buy} > EU^{rent}$ for first-time homebuyers:

$$\frac{y_2^0 + R^h pH_1 - \kappa}{1+r} - pH_1 - \beta y_2^0 + (\phi - \phi_0 + \rho) H_1 \geq 0 \tag{14}$$

To study the impact of interest rate r on renters' c_1, rewrite(12) as

$$c_1 = y_1^0 + \frac{y_2^0 + R^h pH_1 - \kappa}{1+r} - pH_1,$$

Taking derivative, we obtain the marginal effect of monetary easing on type 0 households' consumption

$$\frac{\partial c_1}{\partial r} = -\frac{y_2^0 + R^h pH_1 - \kappa}{(1+r)^2} + \frac{R^h - (1+r)}{1+r}\frac{\partial p}{\partial r} < 0 \tag{15}$$

That is, for type 0 households who decide to purchase houses (i.e., for first-time homebuyers), their period-1 consumption still follows the direction implied by conventional wisdom, that is, an easing of monetary policy increases their consumption.

III Type 1 Households (Existing Homeowners)

A type 1 household chooses between stay or trade up by solving the following problem

$$V^1 = \max\left\{EU^{\text{stay}}, EU^{up}\right\} \tag{16}$$

If he stays, his utility becomes

$$\begin{aligned} EU^{\text{stay}} &= I_{\{c_1\geq 0\}}\cdot(c_1+\beta Ec_2+\phi H_1)+I_{\{c_1<0\}}\cdot(-\infty) \\ \text{s.t.} \quad & c_1 = y_1^1 \\ & c_2 = y_2^1 + R^h\cdot pH_1 - \chi_1 \end{aligned}$$

where χ_1 denotes the legacy mortgage an existing homeowners has. For simplicity, we assume that households do not pay interest on the legacy mortgage. If he chooses to trade up for a larger house, his utility is

$$\begin{aligned} EU^{up} &= I_{\{c_1\geq 0\}}\cdot(c_1+\beta Ec_2+\phi H_2)+I_{\{c_1<0\}}\cdot(-\infty) \\ \text{s.t.} \quad & c_1 + pH_2 + \chi_1 = y_1^1 + m + pH_1 \\ & m = \lambda_1 pH_2 \\ & c_2 = y_2^1 + R^h\cdot pH_2 - m(1+r) - \kappa \end{aligned} \tag{17}$$

The feasibility for Type-1 households to trade-up his existing home is $c_1 \geq 0$. With equation (17) , we can establish the following proposition regarding the feasibility condition for trade up:

Lemma 1 (Trade-up Feasibility Condition): It would be feasible for a type 1 household to trade up when the following *trade-up feasibility* condition holds:

$$y_1^1 \geq \chi_1 - [H_1 - (1-\lambda_1)H_2]p. \tag{18}$$

We assume that $H_1 - (1-\lambda_1)H_2 < 0$ so that the right-hand-side of (18) is always positive. This assumption can be well justified by the fact that in emerging economies, the first house are typically much smaller or has low-quality.

Assiming the feasibility condition (18) is satisfied, it would be optimal for a type 1 household to trade up (i.e. $EU^{up} \geq EU^{\text{stay}}$) when the following condition holds:

$$\beta ER^h p(H_2 - H_1) + (pH_1 + \phi H_2 - \phi H_1) \geq [(1-\lambda_1)pH_2 + \chi_1 + \beta(1+r)\lambda_1 pH_2 - \beta\chi_1] + \kappa \tag{19}$$

Reorganzing (19), we obtain the following *trade-up optimality* condition.

Lemma 2 (Trade-up Optimality Condition): For a given combination of house price, interest rate, and expected return (p, r, ER^h), there exists a threshold of legacy mortgage balance, χ_1, below which it is optimal for the existing homeowner to trade up. Mathematically, an existing household will trade up if his legacy mortage banlance satisfies

$$\chi_1 \leq \overline{\chi} \equiv \frac{[\beta ER^h(H_2 - H_1) + L]p + \phi(H_2 - H_1) - \beta\lambda_1 H_2 p(1+r) - \kappa}{1-\beta} = g(p,r) \tag{20}$$

where $L \equiv H_1 - (1-\lambda_1)H_2$.

The intuition for Lemma 2 is as follows: because existing homeowners need to pay off the legacy mortgage before switching to a larger house, the legacy mortgage is effectively an upfront cost for the existing homeowners. Therefore, the lower it is, the more likely the existing homeowners will find it optimal to trade up.

We would like to investigate the sign of $\frac{\partial g(p,r)}{\partial p}$. Equation (20) implies that

$$\frac{\partial g(p,r)}{\partial p} = \frac{1}{1-\beta}\left[\beta ER^h(H_2 - H_1) + L - \beta\lambda_1 H_2(1+r)\right] \tag{21}$$

$\frac{\partial g(p,r)}{\partial p} > 0$ requires the following condition to hold

$$\beta ER^h(H_2 - H_1) + L - \beta\lambda_1 H_2(1+r) > 0$$

or

$$\left[\beta ER^h - (1-\lambda_1) - \beta\lambda_1(1+r)\right] H_2 > \left(\beta ER^h - 1\right) H_1 \tag{22}$$

The LHS is the net utility benefit of trading up to H_2, which is the utility benefit from the capital gain of trading up minus the utility cost of investing in H_2, whereas the RHS is the net utility cost of holding H_1. Intuitively, there are two *counterveiling* forces for an increase in house price p to affect the tradeup decision: On the one hand, a higher p makes it more costly to buy each square meter of the larger house today (as reflected in the mortgage principal and interest payment term $\beta\lambda_1 H_2 p(1+r)$), which would make type 1 households less willing to trade up; on the other hand, a higher p today also increases the future capital gain (as future house price is a linear function of the current house price), which would make type 1 households more willing to trade up. The net effect depends on the relative strength of these two forces, and we will consider parameterizations under which the second force dominates the first one.

Furthermore, Equation (21) implies that

$$\frac{\partial\left(\frac{\partial g(p,r)}{\partial p}\right)}{\partial ER^h} = \frac{\beta(H_2 - H_1)}{1-\beta} > 0$$

That is, the higher is the anticipated housing returns, the more likely type 1 households would trade up their homes in response to a marginal increase in house prices. This is because the higher is ER^h, the larger is the net utility benefit of trading up when house price increases, thus the more likely is the probability of trading up.

Regarding the impact of interest rate on the threshold of legacy mortgage for housing tradeup, note that

$$\frac{\partial g(p,r)}{\partial r} = -\frac{\beta\lambda_1 H_2 p}{1-\beta} < 0.$$

In other words, a decrease in r reduces the financing cost for trading up, which encourage more type-1 households to trade up, even if house price is unchanged. In this section, we assume $\frac{dp}{dr} < 0$, which we will prove in the next section. We now establish the following proposition on the overall effects of a change in interest rate on the threshold legacy mortgage $\overline{\chi}$.

Proposition 1 (Effects of Monetary Easing along Extensive Margin): Assume both the trade-up feasibility condition (**18**) and (22) hold. A reduction of interest rate makes it more likely for existing homeowners to trade up their homes, by increasing the lower bar of legacy moretgage, i.e.,

$$\frac{d\overline{\chi}}{dr} = \frac{dp}{dr}\frac{\partial g(p,r)}{\partial p} + \frac{\partial g(p,r)}{\partial r} < 0. \tag{23}$$

Moreover, a higher expected returns for housing purchase makes it more likely for type 1 households to trade up their homes in response to a marginal reduction in interest rates.

$$\frac{\partial\left(d\overline{\chi}/dr\right)}{\partial ER^h} < 0. \tag{24}$$

The intution for Proposition 1 is as follows: a reduction in interest rates makes tradeup more likely to happen for two reasons. The first is a direct effect, by reducing the future interest payment associated with purchasing a larger house. The second is an indirect effect via equilibrium house prices: by increasing housing prices, and thus the capital gains, a reduction of interest rates increases the marginal benefit of trading up. Moreover, a higher expected return for housing would increase the response of probability of tradeup to a reduction of interest rate via the indirect effect.

Now we study the impact of a change in interest rate on non-housing consumption of type 1

household who decides to trade up. Note that

$$\begin{aligned} c_1^1 &= y_1^1 + pH_1 + \lambda_1 pH_2 - pH_2 - \chi_1 \\ &= y_1^1 - \chi_1 + [H_1 - (1-\lambda_1)H_2]p \end{aligned} \tag{25}$$

Taking derivative of (25), we can establish the following proposition:

Proposition 2 (Effects of Monetary Easing along Intensive Margin): Assume both the trade-up feasibility condition (**18**) and (22) hold. A reduction of interest rates reduces the non-housing consumption of an individual type 1 household who decides to trade up his/her house. In other words,

$$\frac{dc_1^{up}}{dr} = [H_1 - (1-\lambda_1)H_2]\frac{dp}{dr} > 0 \tag{26}$$

where the superscript "up" denotes those type 1 households who trade up their homes. The inequality follows because $L \equiv H_1 - (1-\lambda_1)H_2 < 0$.

The intuition is as follows. In this model, there are two forces for interest rate to affect period 1 consumption. 1) Wealth effect: As p increases following monetary easing, the wealth of existing homeowners increases, enabling the trade-up household to consume more in period 1. 2) Investment cost effect: However, a higher p also increases the per unit cost of buying a larger house today and the associated downpayment cost $(1-\lambda_1)H_2 p$. Under the assumption of $H_1 - (1-\lambda_1)H_2 < 0$, the wealth effect is dominated by the investment cost effect, ultimately decreasing the period-1 consumption of the households who decide to trade up in equilibrium. Then why do they want to do so? Because they expect a much higher house price in the second period and therefore a higher life-time consumption.