Irving Fisher Committee on
Central Bank Statistics

BIS

12th biennial IFC Conference: "Statistics and beyond: new data for decision making in central banks"

22-23 August 2024

Bi-dimensional sentiment indicators based on new generative AI capabilities – a pilot study in the housing market in Israel[1]

Alon Bartal, Michael Gurkov, Ariel Mantzura, Yulia Nudelman,
Darin Vaisman and Inbal Yahav Shenberger,
Bank of Israel

[1] This contribution was prepared for the conference. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

Bi-dimensional sentiment indicators based on new generative AI capabilities

A pilot study in the housing market in Israel

Alon Bartal, Michael Gurkov, Ariel Mantzura, Yulia Nudelman, Darin Vaisman, Inbal Yahav Shenberger; Bank of Israel

Abstract

This study explores the innovative application of Large Language Models (LLMs) to gauge public sentiment in the real estate market, leveraging unstructured text data to uncover insights typically inaccessible through traditional analytical methods. Our research introduces a novel 'two-dimensional sentiment' framework that combines entity recognition with sentiment analysis to map the sentiment landscape across key stakeholders and topics within the real estate domain. This approach not only enables the identification of main stakeholders and pivotal topics but also facilitates the extraction of sentiment for each stakeholder-topic pairing, providing a robust and nuanced understanding of market sentiments. This kind of analysis is particularly valuable in economic research, where understanding the nuances of market sentiment can offer deep insights.

Contents

Introduction 3

Data 4

Methodology 4

Workflow ....

Few shot prompting 6

Promp structure 6

Findings 7

Sentiment heatmaps 7

Correlations with relevant indicators 8

Conclusions 10

References 11

Apendix 12

1. Introduction

Sentiment analysis, also known as opinion mining (Dave et al., 2003), has gained popularity in recent years due to the advent of new technological capabilities that enable the processing of large amounts of data in various forms, such as textual data from different media sources. Alongside this is the rapid evolution of new NLP methods, with particular importance given to large language models (LLM) trained on vast amounts of data, which are suitable for efficiently performing tasks such as sentiment analysis.

Monitoring public sentiment and understanding the underlying forces in various sectors of economic activity is crucial for informed policymaking, as economic agents make decisions based, among other factors, on opinions, attitudes, and emotions toward different issues (Akerlof et al., 2010). Moreover, understanding public sentiment towards economic matters is vital for central bank policymaking, as it can help policymakers gauge the effectiveness and acceptability of their decisions. For instance, if economic agents are optimistic about the housing market, they might be more likely to invest in property (Soo, 2018) or take on mortgages, influencing overall economic activity. Conversely, negative sentiment might lead to reduced housing market activity and a potential economic slowdown. By considering public sentiment, central banks can better craft policies to support a stable and balanced housing market, thereby supporting financial stability.

In this work, which is part of a broader strategic project at the Bank of Israel, we attempt to gauge public sentiment held by important economic agents regarding different economic topics as reflected in economic news articles. Much has been written on the transmission channel of public sentiment via news articles, or in other words, how news articles influence public sentiment and vice versa (Tetlock, 2007; Shiller, 2017; Soo, 2018). For this end, we harness the approach of aspect-based sentiment analysis. Aspect-based sentiment analysis is a sub-field of sentiment analysis that focuses on identifying and extracting sentiment towards specific aspects (topics) in a given text. We add to this approach the identification of important entities that sentiment is ascribed to and the extraction of these entities' sentiments towards various topics or subjects. In this way, we aim to mitigate possible biases present when performing general polarity sentiment classification for each document which represents multiple agents sentiment including even the journalists and editors one (Soo, 2018) hence representing contradicting sentiments.

Specifically, we identified important entities (stakeholders) in the housing market and significant topics. For each pair, we provide a polar sentiment score (positive, neutral, or negative) using GPT[1] and appropriate prompting. Many studies (See

[1] We used in this work GPT4o mini.

Keiri et al, 2023; Krugman et al., 2024) have examined the advantages of sentiment analysis via GPT relative to suitable BERT models or lexicon-based models and found that GPT was superior in zero-shot and few-shot scenarios. Our results are displayed in bi-dimensional heat maps where each cell represents the sentiment for a specific pair (stakeholder, topic). For each pair, we can construct a dynamic time series to examine correlations with appropriate indicators. We find that performing sentiment analysis with GPT is considerably well aligned with human-based sentiment analysis and that the time series constructed show relatively strong correlations with relevant indicators.

By understanding these correlations and the sentiments of key stakeholders, policymakers and economic analysts can make more informed decisions that better reflect public sentiment, leading to more effective and accepted policies. This can ultimately result in a more stable and balanced economic environment, benefiting both policymakers and the public.

The rest of this work is organized as follows: in Section 2, we describe the data used in this pilot study; in Section 3, we explain the methodology; in Section 4, we present the findings; and in Section 5, we conclude.

2. Data

The data utilized in this study comprises newspaper articles relevant to the real estate market, provided by Ifat Media, a commercial data provider. The initial dataset includes approximately 100,000 articles. To ensure the analysis focuses on substantial content, we filtered the articles to include only those with a minimum length of 200 characters. This filtering process resulted in a refined dataset of about 70,000 articles.

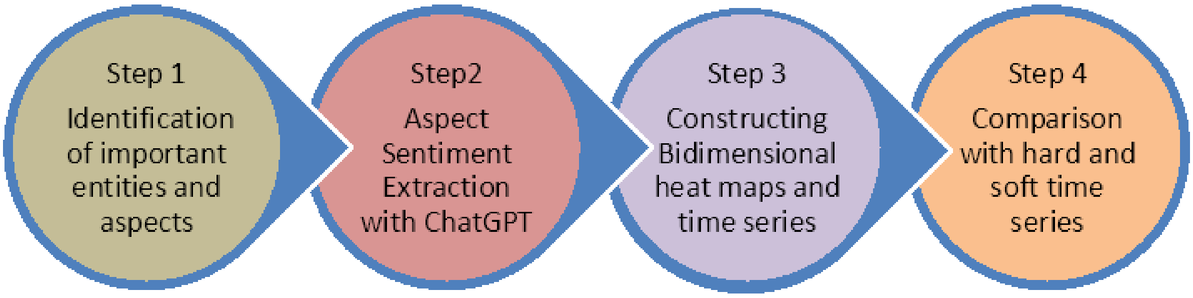
The articles span a significant time period from December 26, 2006, to June 24, 2024, offering a comprehensive view of the real estate market over 17 years. The dataset is characterized by daily frequency, providing a granular temporal resolution for sentiment analysis.

3. Methodology

The methodology employed in this study leverages aspect-based sentiment analysis, which focuses on identifying and extracting sentiments towards specific aspects within the text. This approach is enhanced by identifying important entities and extracting their sentiments towards various topics or subjects. The innovative aspect of this methodology lies in its ability to pinpoint the sentiments of key stakeholders in the Israeli real estate industry regarding particular topics, providing a nuanced understanding of the market's sentiment landscape.

Workflow

Our methodology consists of the following steps:



Step1: Identification of important entities and aspects

The methodology involves identifying significant stakeholders in the real estate market and extracting their sentiments towards relevant topics. The stakeholders considered in this study include construction companies, investors, mortgage borrowers, first-time homebuyers, home upgraders, and home renters. The topics of interest are housing market demand, housing market supply, housing market policies, housing market construction starts, rental prices, rental housing, mortgages, and interest rates.

Step2: Aspect Sentiment Extraction with GPT

The sentiment extraction process utilizes Large Language Models (LLMs) with a specific prompt designed to guide the model in evaluating sentiments. The prompt is structured to ensure clarity and precision in the sentiment analysis task. The process employs few-shot prompting, a technique that provides the model with a few examples to demonstrate the desired task, enhancing its ability to understand and perform the task accurately. Aspect-based sentiment analysis (ABSA) is a sub-field of sentiment analysis aimed at determining the sentiment expressed about specific aspects within a given text. In this study, ABSA is extended to identify and assess the sentiments of important stakeholders in the Israeli real estate market towards specific topics. This allows for a detailed examination of how different entities feel about various market factors.

Step 3: Constructing Bi-dimensional heat maps and time series

To visualize the sentiment analysis results, we constructed bi-dimensional heat maps where each cell represents the sentiment for a specific pair (stakeholder, topic). These heat maps provide an intuitive way to observe the distribution and intensity of sentiments across different stakeholders and topics.
For each pair, we also constructed dynamic time series to examine the evolution of sentiments over time. This involved aggregating the sentiment scores for each pair on a daily basis and plotting them to observe trends, fluctuations, and correlations with real-world events or economic indicators.

Step 4: Comparison with relevant indicators

We compared the sentiment scores with relevant economic indicators to validate the alignment and predictive power of the sentiment analysis. For instance, we examined correlations with construction starts, housing market supply, and other key indicators to assess how well the sentiments captured by our model align with actual economic activities.

Few Shot Prompting

Few-shot prompting is demonstrated in the system part, where both correct and incorrect examples are provided. This technique helps the model understand the task by showing examples of acceptable and unacceptable responses. This approach enhances the model's ability to accurately perform the sentiment analysis task.
This methodology, combining aspect-based sentiment analysis with entity-specific sentiment extraction, provides a detailed and nuanced understanding of the sentiments expressed by key stakeholders in the Israeli real estate market. The use of a structured prompt with few-shot examples ensures accurate and consistent sentiment evaluation, making this approach both innovative and effective for the study at hand.

Prompt Structure

The prompt consists of two main parts: the system part and the user part.

System Part:

```text
You are a real estate expert. Your goal is to assess the sentiment of an important stakeholder in the Israeli real estate industry regarding a certain topic in a supervised fashion (from news articles). I'll provide you with the category of the stakeholder and the topic of interest. You need to evaluate the sentiment of the stakeholder with respect to the topic as either positive, negative, or neutral.
Please return the result in one-word format with no additional text or explanation.
If the stakeholder with respect to the topic is not mentioned in the article return irrelevant.

Response requirements:

Incorrect Example: 'Here is the evaluated sentiment: positive.'
Correct Example: 'positive'.
Incorrect Example: 'negative. The article discusses the current state ...'
Correct Example: 'negative'.
```

The response should be in English only (not Hebrew!!).

Incorrect Example: 'נייטרלי'

Correct Example: 'neutral'.
Correct Example: 'irrelevant'.
```

### User Part:

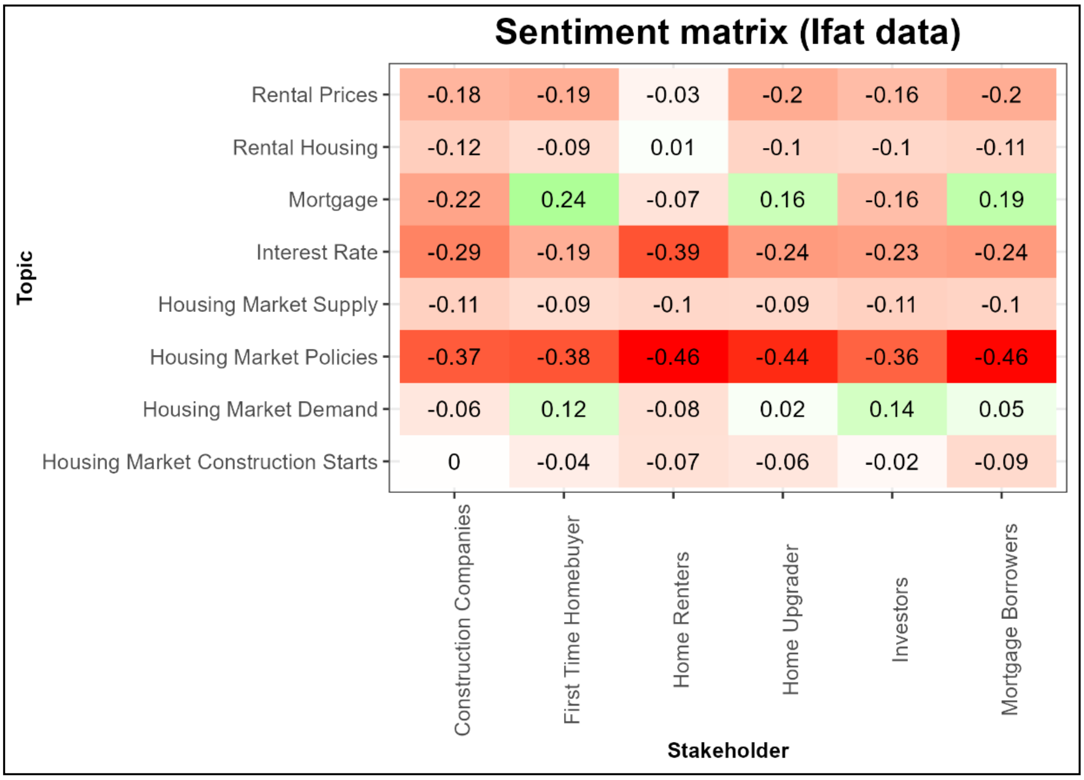
```text
Below is the text of a news article.
The stakeholder's category is: """ + "".join(stakeholder) + \
""" The topic is: """ + "".join(topic) + \
""" Article's text:
```

Main Components of the Prompt

Stakeholder and Topic Identification: The system part specifies the category of the stakeholder and the topic of interest, ensuring the model focuses on relevant aspects.
Sentiment Evaluation: The model is instructed to evaluate the sentiment as either positive, negative, or neutral. If the stakeholder and topic are not mentioned, the model should return "irrelevant."
Response Format: The response is required to be in a one-word format (positive, negative, neutral, or irrelevant) with no additional text or explanation. This ensures consistency and clarity in the output.
Language Specification: The model is instructed to provide responses in English only, avoiding any potential confusion with Hebrew terms.

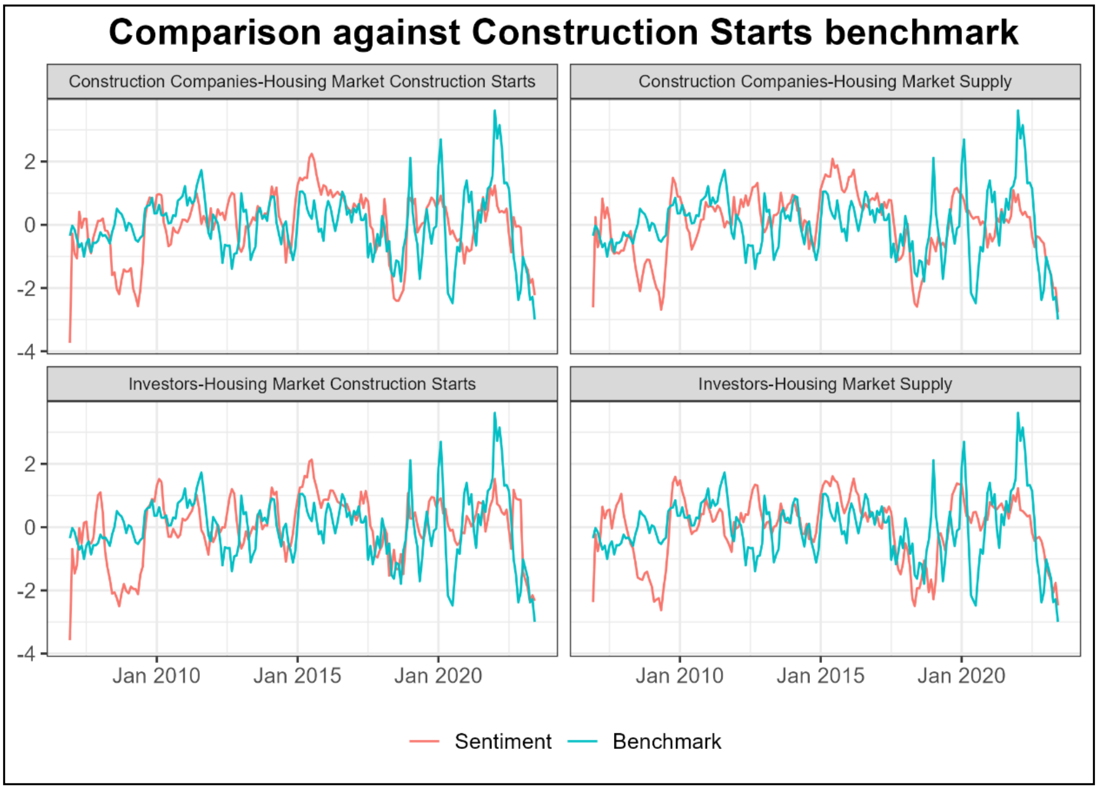
4. Findings

Sentiment Heatmaps
```

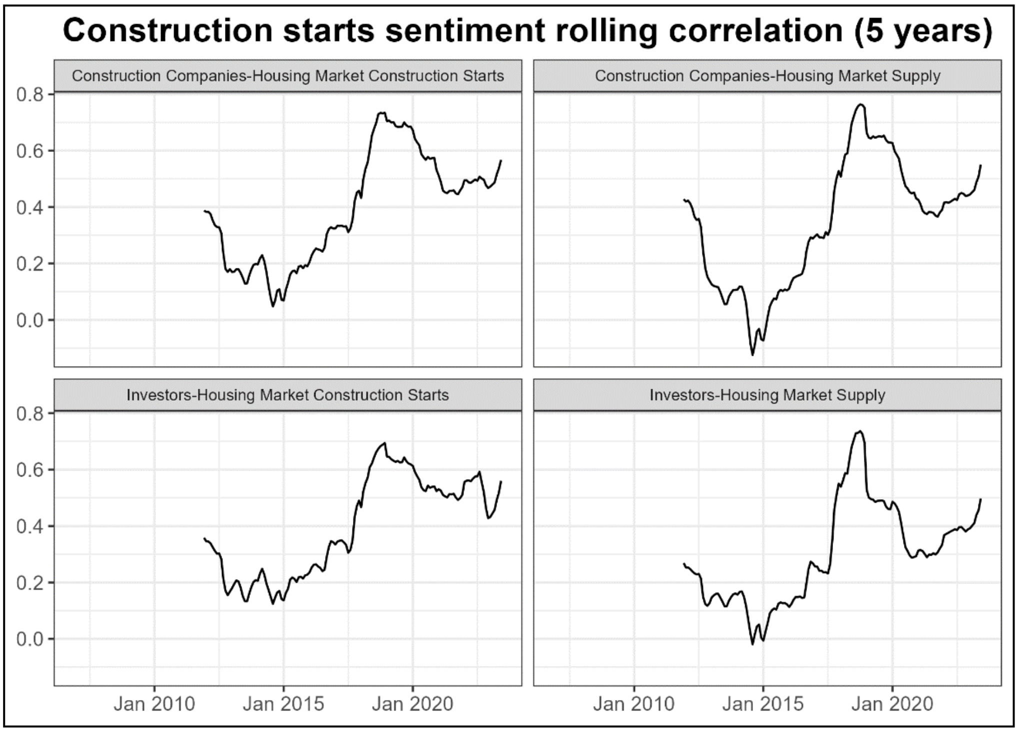


Heat map of average sentiment for the year 2013. We can see slightly positive sentiments with respect to mortgage among first-time home buyers and home upgraders, while home renters and mortgage borrowers are characterized by a rather negative sentiment towards housing market policies.

Correlations with relevant indicators

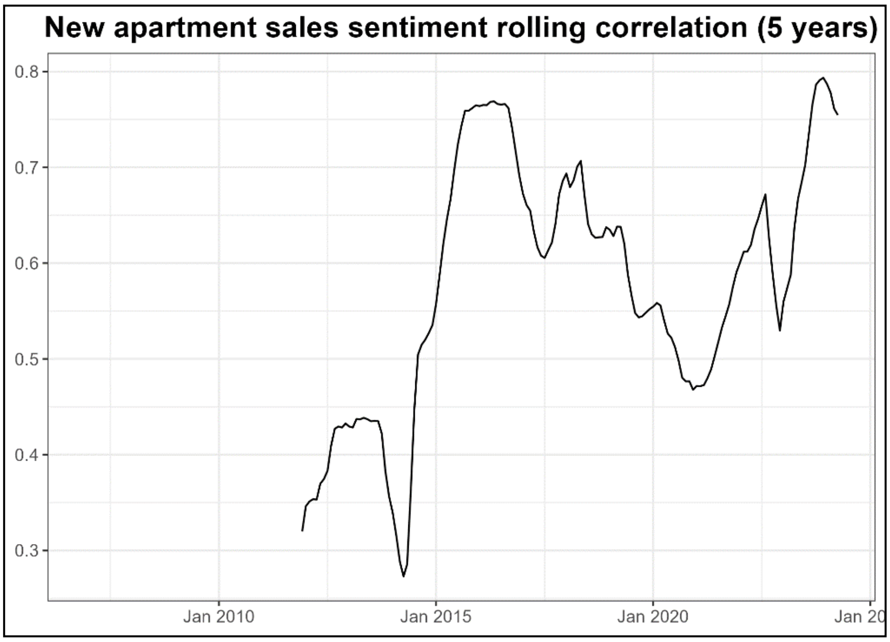


The comparison of extracted sentiment against economic indicators (construction starts) reveals that while sentiment often follows the benchmark trends, there are notable deviations the benchmark seems to be characterized by larger fluctuations relative to the sentiment series. In the sentiment categories of Construction Companies-Housing Market Construction Starts and Supply, as well as Investors-Housing Market Construction Starts and Supply, sentiment generally aligns with the benchmark but shows significant divergence at certain points. Overall, the relationship between sentiment and benchmarks varies, indicating that sentiment analysis can be insightful but is subject to fluctuations due to other influences.



The five-year rolling correlation between sentiment and economic benchmarks generally trends upward over time. Initially, correlations are low or even negative, indicating a weak or inverse relationship. However, from around 2015 onwards, the correlations increase significantly, reaching levels above 0.7 before declining slightly, though they remain relatively high. This trend suggests that sentiment has increasingly aligned with economic indicators over the years, reflecting a strengthening relationship between sentiment and benchmarks in the housing market.

The comparison against the new apartment sales indicator shows that sentiment fluctuations are generally of a larger magnitude than those of the benchmark. While sentiment and benchmark values are mostly aligned, there are periods of notable divergence, indicating potential shifts in underlying factors. Around 2019, the sentiment series exhibits a sharp decline followed by a quick recovery, while a similar dip appears in the benchmark series, albeit slightly later. This pattern suggests a possible leading/lagging relationship between sentiment and the benchmark. Overall, the relationship between sentiment and the benchmark varies over time, reflecting the complexity of the factors at play.



Consistent with previously observed patterns, the five-year rolling correlation rises, peaking above 0.7 by 2015, which suggests a strong alignment between sentiment and new apartment sales during this period. This is followed by a declining trend, which subsequently reverses into a sharp increase, reflecting fluctuating yet generally strong correlations over time. This trend underscores that while sentiment has increasingly aligned with new apartment sales, external factors can cause temporary disruptions in this relationship..

5. Conclusion

In conclusion, our pilot study demonstrates the potential of using aspect-based sentiment analysis combined with advanced language models like GPT to gauge public sentiment in the housing market. The ability to extract and analyse sentiments of specific stakeholders towards various topics provides a detailed and nuanced understanding of market dynamics. By aligning these sentiments with relevant economic indicators, policymakers and analysts can make more informed decisions, potentially leading to more effective and accepted policies.

Our findings suggest that sentiment analysis using GPT correlates considerably well with relevant economic indicators over time. The dynamic time series and bi-dimensional heat maps provide a comprehensive view of how sentiments evolve and their potential impact on economic activities. This approach can help central banks and policymakers better understand public sentiment, enabling them to craft

policies that support a stable and balanced housing market and, ultimately, financial stability.
Future research could expand this methodology to other sectors and incorporate additional data sources to further validate and refine the approach. The integration of sentiment analysis into economic modelling and forecasting holds promise for enhancing the understanding of economic behaviours and improving policy decisions.

6. References

Akerlof, G.A. and Shiller, R.J., 2010. Animal spirits: How human psychology drives the economy, and why it matters for global capitalism. Princeton university press.
Shiller, R.J., 2017. Narrative economics. American economic review, 107(4), pp.967-1004.

Dave, K., Lawrence, S. and Pennock, D.M., 2003, May. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web (pp. 519-528).

Kheiri, K. and Karimi, H., 2023. Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning. *arXiv preprint arXiv:2307.10234*.

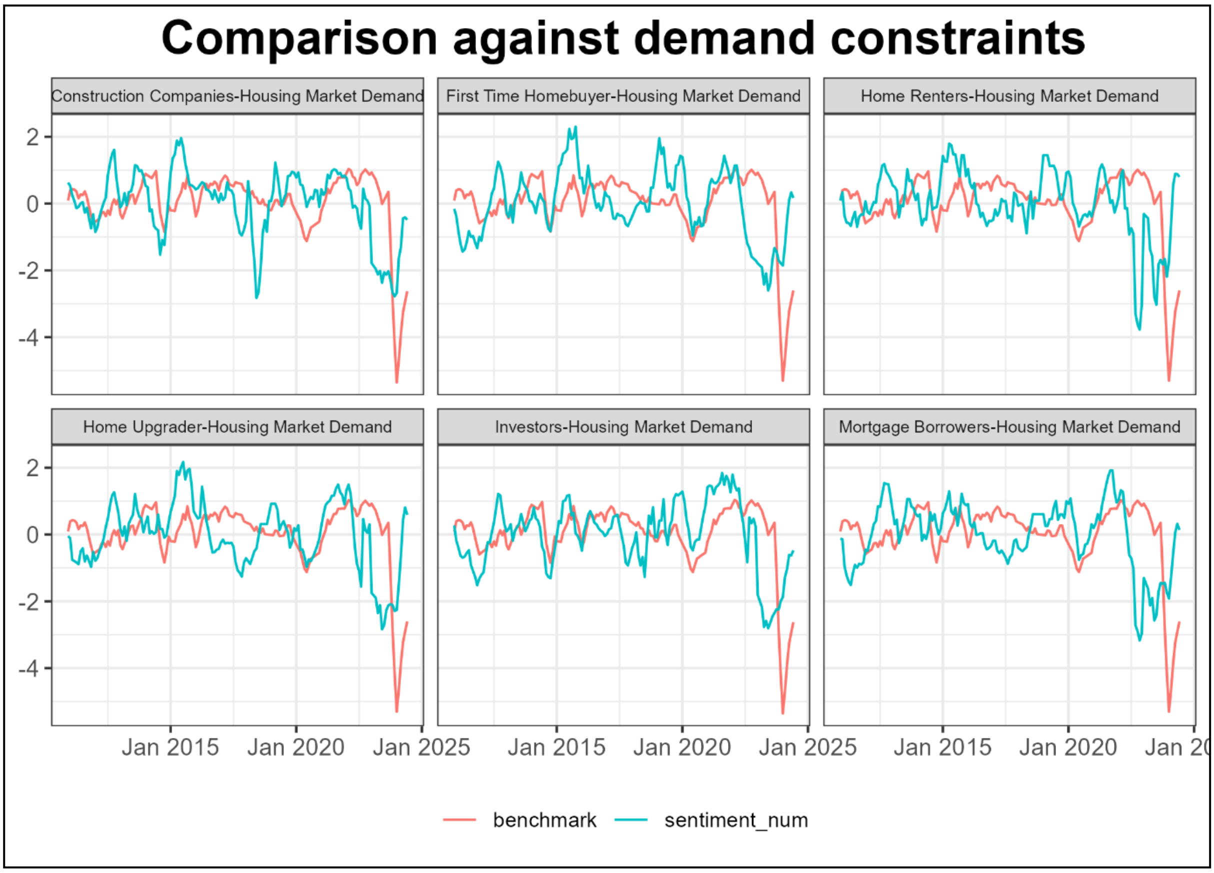
Krugmann, J.O., Hartmann, J. Sentiment Analysis in the Age of Generative AI. *Cust. Need. and Solut.* **11**, 3 (2024). https://doi.org/10.1007/s40547-024-00143-4

Soo, C.K., 2018. Quantifying sentiment with news media across local housing markets. The Review of Financial Studies, 31(10), pp.3689-3719.

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

7. Appendix

In extending the analysis of sentiment, we have introduced a comparison to “soft indicators,” which are derived from survey data and align more closely in nature with sentiment than the previously utilized hard economic indicators. While this approach offers promising insights, the results should be viewed as preliminary. One key challenge is establishing a clear correspondence between specific survey questions or topics and the sentiment aspects analysed. Survey responses often capture broad or nuanced opinions that may not directly align with the sentiment dimensions extracted, making it difficult to create a straightforward mapping between them. As such, while initial findings suggest a potential relationship, further refinement in linking survey content with sentiment aspects is necessary to draw more robust conclusions.



Comparison against demand constraints
Construction Companies-Housing Market Demand
First Time Homebuyer-Housing Market Demand
Home Renters-Housing Market Demand
Home Upgrader-Housing Market Demand
Investors-Housing Market Demand
Mortgage Borrowers-Housing Market Demand
2
0
-2
-4
Jan 2015
Jan 2020
Jan 2025
benchmark
sentiment_num

BI –DIMENSIONAL SENTIMENT INDICATORS BASED ON NEW GENERATIVE AI CAPABILITIES – A PILOT STUDY IN THE HOUSING MARKET

Alon Bartal, Michael Gurkov, Ariel Mantzura, Yulia Nudelman, Darin Vaisman, Inbal Yahav Shenberger

SENTIMENT FROM NEWS ARTICLES

Stakeholder

on

Topic

Sentiment

- ✓ 18 articles
- ✓ ~36,000 articles
- ~100,000 articles

- ✓ GPT 3.5
- GPT 4o
- Prompt?

- ✓ Human evaluation
- ✓ Economic indicators
- Soft indicators

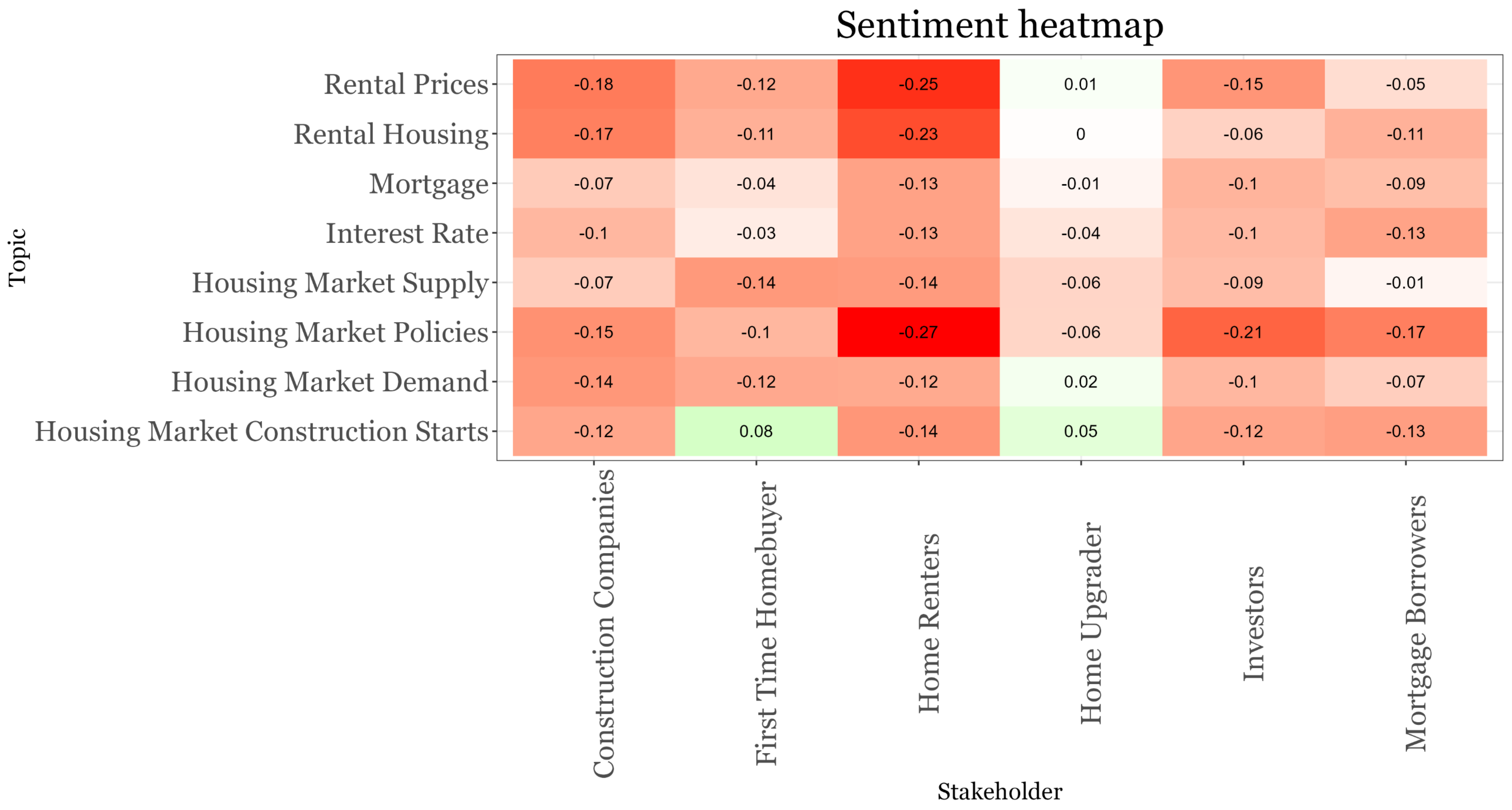
ASPECT SENTIMENT ANALYSIS

Stakeholders

1. Construction companies
2. Investors
3. Mortgage borrowers
4. First time homebuyer
5. Home upgrader
6. Home renters

Topics

1. Housing market demand
2. Housing market supply
3. Housing market policies
4. Housing market construction starts
5. Rental prices
6. Rental housing
7. Mortgage
8. Interest rate


Sentiment heatmap
Topic
Rental Prices
Rental Housing
Mortgage
Interest Rate
Housing Market Supply
Housing Market Policies
Housing Market Demand
Housing Market Construction Starts
-0.18
-0.12
-0.25
0.01
-0.15
-0.05
-0.17
-0.11
-0.23
0
-0.06
-0.11
-0.07
-0.04
-0.13
-0.01
-0.1
-0.09
-0.1
-0.03
-0.13
-0.04
-0.1
-0.13
-0.07
-0.14
-0.14
-0.06
-0.09
-0.01
-0.15
-0.1
-0.27
-0.06
-0.21
-0.17
-0.14
-0.12
-0.12
0.02
-0.1
-0.07
-0.12
0.08
-0.14
0.05
-0.12
-0.13
Construction Companies
First Time Homebuyer
Home Renters
Home Upgrader
Investors
Mortgage Borrowers
Stakeholder

PROMPT – SYSTEM PART(INSTRUCTION)

You are a real estate expert. Your goal is to assess the sentiment of an important stakeholder in the Israeli real estate industry regarding a certain topic in a supervised fashion (from news articles). I'll provide you with the category of the stakeholder and the topic of interest. You need to evaluate the sentiment of the stakeholder with respect to the topic as either positive, negative, or neutral.

Please return the result in one-word format with no additional text or explanation.

If the stakeholder with respect to the topic is not mentioned in the article return irrelevant.

PROMPT – FEW SHOT APPROACH

Response requirements:

Incorrect Example: 'Here is the evaluated sentiment: positive.'

Correct Example: 'positive'.

Incorrect Example: 'negative. The article discusses the current state ...'

Correct Example: 'negative'.

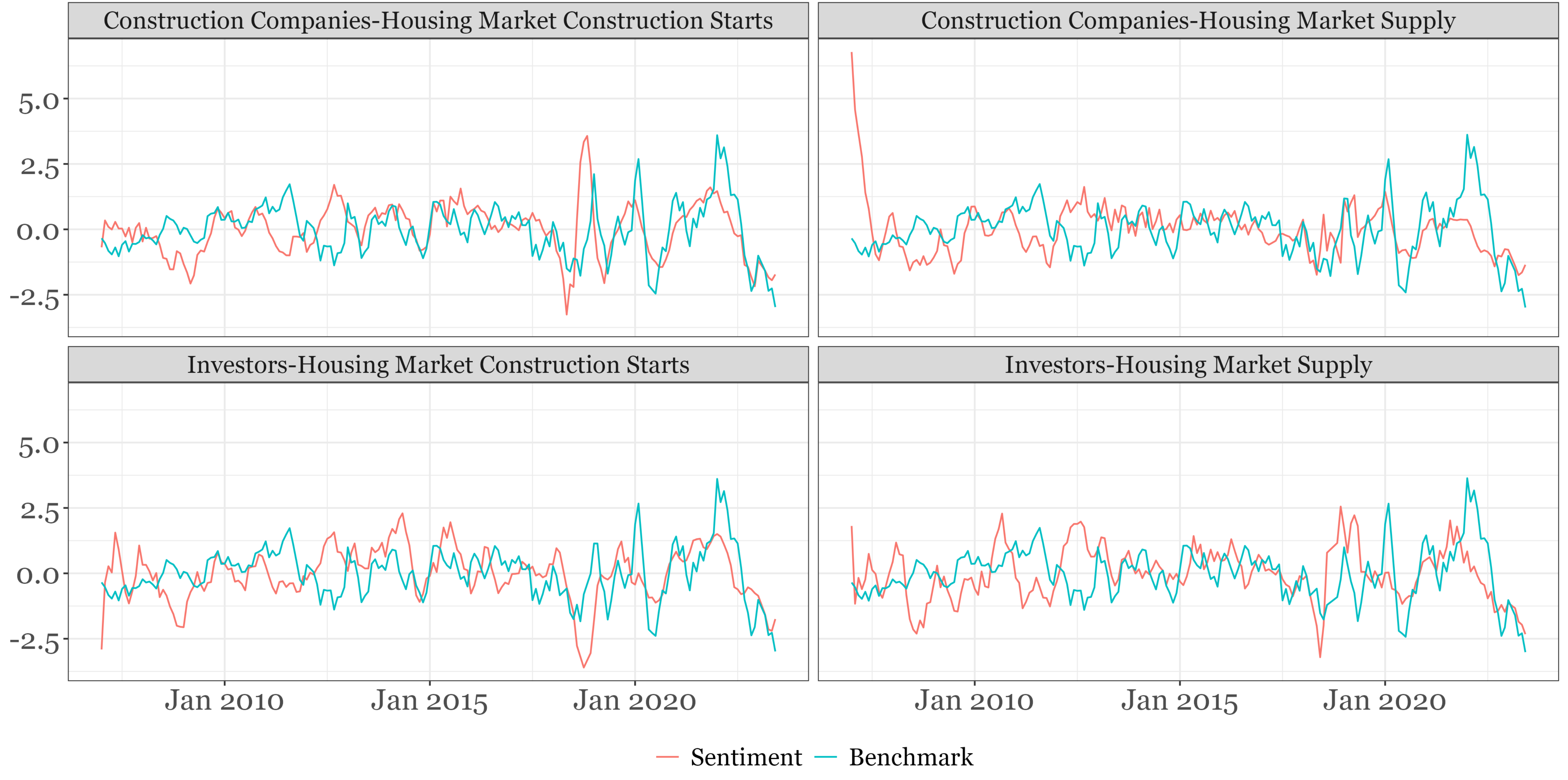
The response should be in English only (not Hebrew!!).

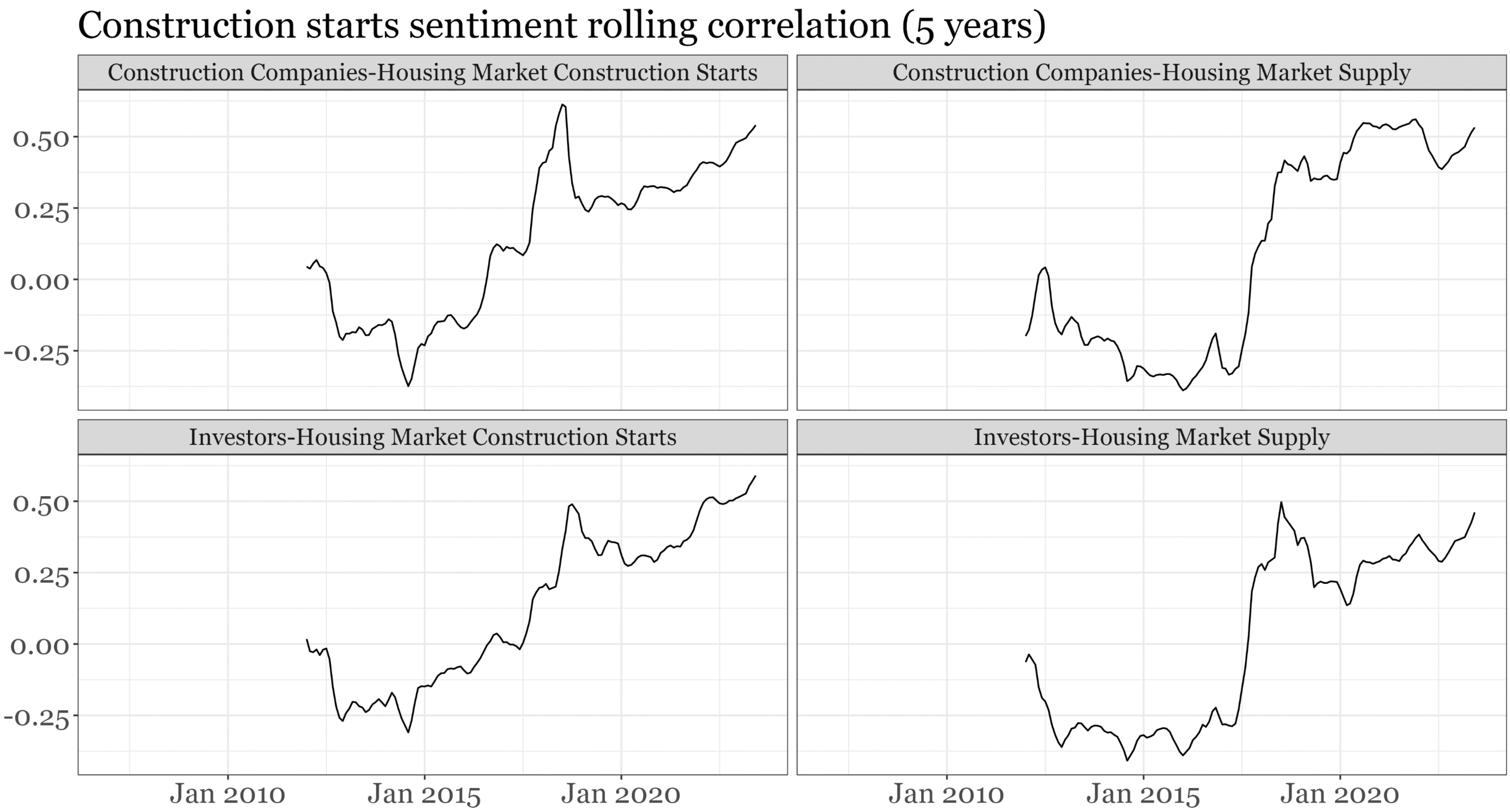
Incorrect Example: "נייטרלי"

Correct Example: 'neutral'.

Correct Example: 'irrelevant'.

Comparison against Construction Starts benchmark




Construction starts sentiment rolling correlation (5 years)
Construction Companies-Housing Market Construction Starts
Construction Companies-Housing Market Supply
Investors-Housing Market Construction Starts
Investors-Housing Market Supply
0.50
0.25
0.00
-0.25
0.50
0.25
0.00
-0.25
Jan 2010
Jan 2015
Jan 2020
Jan 2010
Jan 2015
Jan 2020

THANK YOU FOR YOUR ATTENTION