When LLMs Go Abroad: Foreign Bias in AI Financial Predictions

S. Cao, C. C. Y. Wang, and X. Yi

Discussion by Bohui Zhang
CUHK-Shenzhen

UMD/SMU/UBS Quant Investment Forum: AI and Finance 2026
Singapore

Summary

Objective: this paper documents "foreign bias" in AI-generated financial analysis

Sample: forecasts generated by ChatGPT and DeepSeek on June 30, 2024, for Chinese listed firms

Methods: an LLM-based forecasting framework combined with experimental tests of competing explanations for forecast divergence

Findings:

- ChatGPT generates systematically more optimistic forecasts for Chinese firms than DeepSeek
- Forecast divergence is associated with cross-country differences in financial news coverage
- Providing the same set of firm-specific Chinese news substantially reduces the forecast gap
- The evidence suggests that differences in effective information availability contribute to AI-generated forecast divergence

My direct feedback

- I like papers on home bias
- Great effort on an intriguing research question
- Smart empirical design
- The evidence documented in this study is convincing

Comments

1. The role of LLMs in financial markets
2. The accuracy of LLMs' forecasts
3. Information intermediaries: LLMs vs. analysts
4. Alternative explanation: investor interaction
5. Traditional home bias
6. Understanding the details of news gap
7. The comparative advantage of ChatGPT
8. Synthetic firms
9. Robustness: forecasts vs. luck

1. The role of LLMs in financial markets

The paper examines LLM-generated stock price forecasts and documents systematic differences across models.

LLMs serve as **virtual analysts** rather than information-processing tools.

Most existing applications of LLMs in finance appear to focus on:

- Information retrieval and summarization
  e.g., 10-K, earnings-call transcripts, analyst reports
- Customer-facing financial services
  e.g., retail investors use LLMs for information search, analysis, and interpretation (Ecker et al., 2026)

In these settings, LLMs primarily **augment human decision-making,** while humans remain responsible for the final investment decision.

1. The role of LLMs in financial markets

This raises a broader question:
Should LLMs be used primarily as forecast generators?

Direct point forecasts may face additional challenges:

- hallucination;
- opaque information sources;
- limited interpretability of forecast outputs.

In contrast, information-processing applications allow users to retain judgment over the final investment decision.

Do investors value AI-generated forecasts, or the information-processing capabilities that support human judgment?

2. The accuracy of LLMs' forecasts

Are LLMs' forecasts accurate enough?

the means of *STOCK_PRICE_ERROR* and *EPS_ERROR*

The paper does not report the summary statistics of *STOCK_PRICE_ERROR* and *EPS_ERROR*. I try to calculate the means implied based on the results of Table 3.

The coefficient of *US_AI* is 1.0863, approximately 12.3% larger relative to DeepSeek's average price prediction errors.

DeepSeek's average **price prediction errors**=1.0863/12.3%=8.83,

42.5% of the mean of stock price, 20.76

57.6% of the median of stock price, 15.53

Are these AI forecasts **reliable**?

3. Information intermediaries: LLMs vs. analysts

What is the uniqueness of LLMs as an information intermediary?

- Similar patterns for financial analysts facing language barriers or local media access.

ELSEVIER

Journal of Financial Economics

Volume 88, Issue 3, June 2008, Pages 581-606

Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts ☆

Kee-Hong Bae [a], René M. Stulz [b], Hongping Tan [c]

This paper examines whether analysts resident in a country make more precise earnings forecasts for firms in that country than non-resident analysts. Using a sample of 32 countries, they find an economically and statistically significant local analyst advantage even after controlling for firm and analyst characteristics.

3. Information intermediaries: LLMs vs. analysts

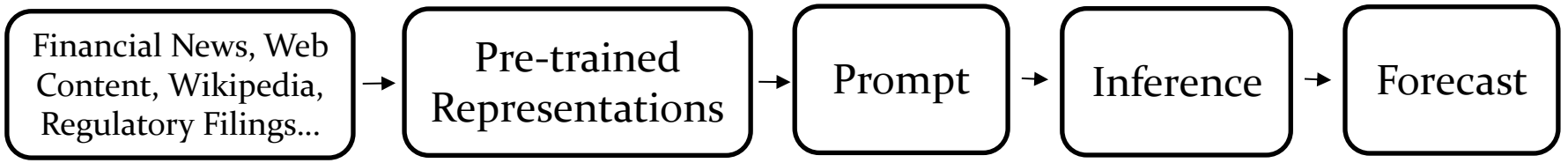
What is the uniqueness of LLMs as an information intermediary?

- The differences between LLMs and financial analysts

Traditional Financial Analyst:

Information is **actively** acquired, filtered, and interpreted by domain experts.

LLM-Based Forecasting:



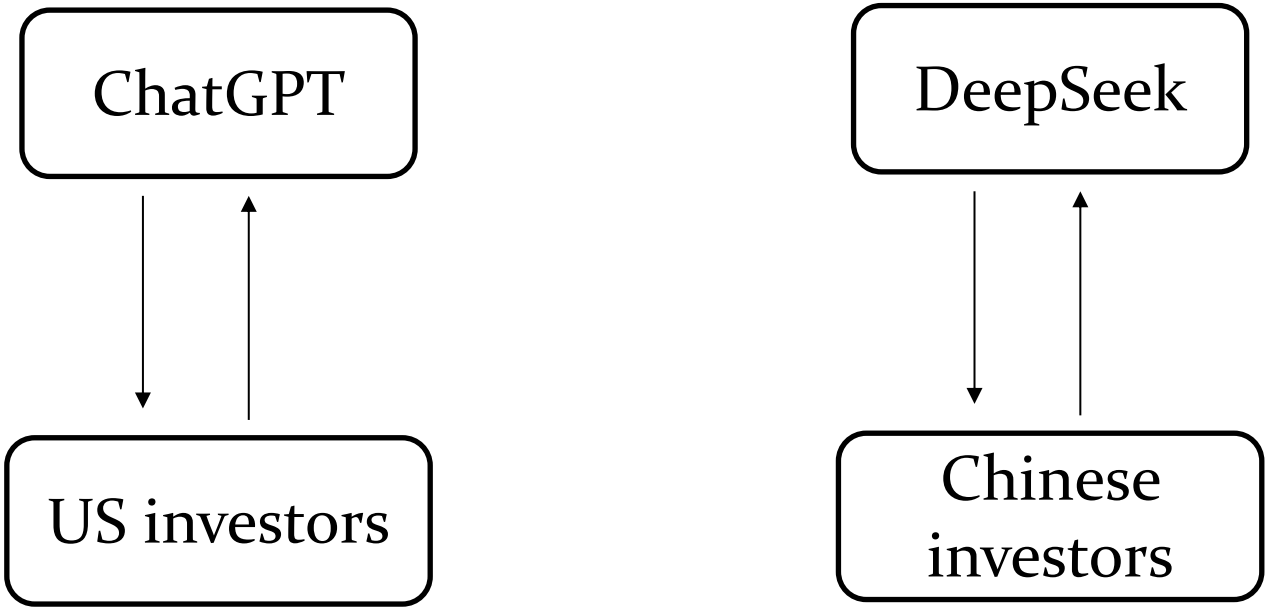
Information is **passively** embedded during training and elicited through prompting.

- **AI–Human forecast gaps** should vary systematically with the importance of soft information, local context, or industry-specific experience
- **LLM gaps** may rely more heavily on public textual information, **while analyst gaps** may be more sensitive to private information channels

4. Alternative explanation: investor interaction

Given the differences in information-production processes, an alternative explanation may be:

- **Model calibration**: ChatGPT is less frequently exposed to Chinese user interactions than DeepSeek.



5. Traditional home bias

The paper frames ChatGPT's optimism as "**foreign bias**"

- The terminology is intuitive because it speaks to the AI bias

Traditional home bias is typically studied as:

- The same investor → different assets (local vs foreign)

A comparison more closely aligned with the traditional home-bias literature would be:

- The same LLM → Chinese firms vs. U.S. firms

ChatGPT or DeepSeek samples:

$$Y_i = \alpha + \beta\, Foreignfirm_i + \gamma' X_i + \delta_j + \varepsilon_i$$

Hypothesis: $\beta > 0$

6. Understanding the details of news gap

The paper interprets news coverage differences as differences in information environments embedded in LLM training corpora.

Cross-country news gaps may arise because:

- Media cover **different firms**.
- Media provide **different information** about the same firm.
- Media interpret the same **information differe**ntly

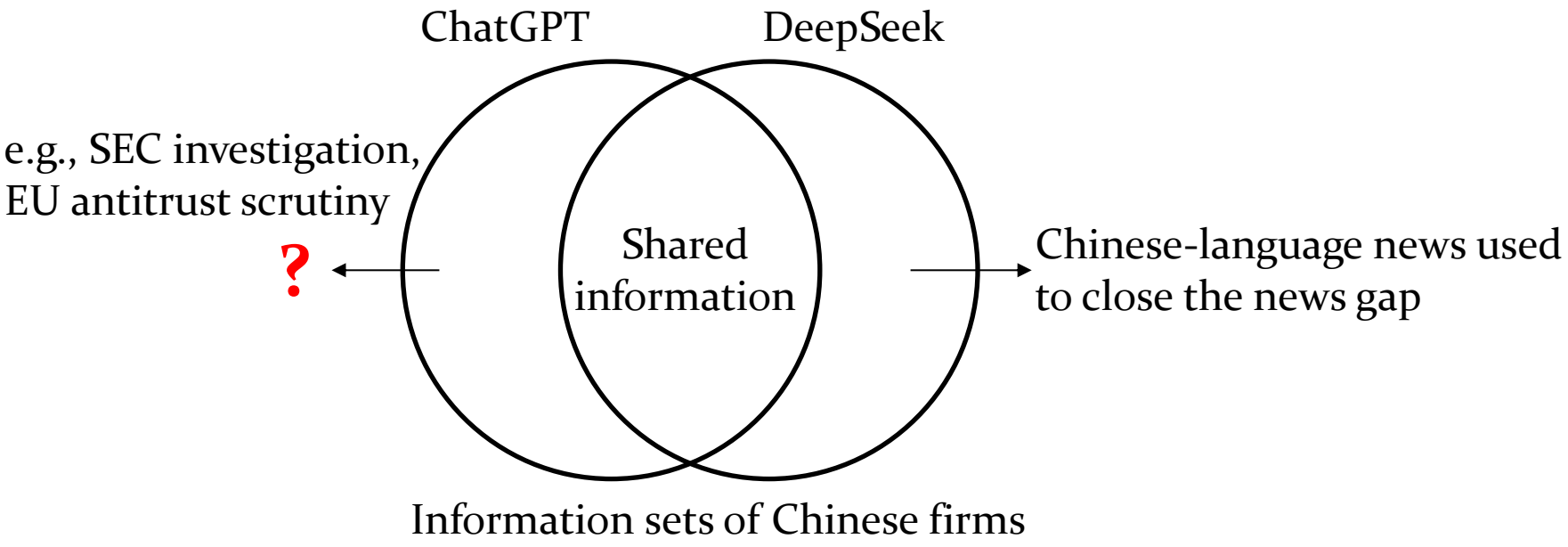
Which of these channels is most closely related to the **information gap** embedded in LLMs?

7. The comparative advantage of ChatGPT

The **news-injection experiment** is convincing: providing firm-specific Chinese-language news to both models substantially reduces the forecast gap.

→ChatGPT lacks some China-specific information.

Does ChatGPT have some advantage over DeepSeek on some scenarios? Eg. among a group of Chinese firms sanctioned by the US government



8. Synthetic firms

The paper tests function-centric explanations using **100 synthetic Chinese firm names** without corresponding real-world entities or media coverage.

Some thoughts:

- If synthetic firms have no retrievable firm-specific information, why can LLMs still **generate meaningful forecasts**?
- The synthetic-firm evidence appears to suggest that forecasts may not rely on firm-specific information.
- Does this create some tension with the information-gap interpretation?
- How can synthetic firms have financial information?

9. Robustness: forecasts vs. luck

a. The results are generated based on prediction by the end of June, 2024.

- Could you try predictions based on different periods?
- To repeatedly elicit forecasts over time and examine how ChatGPT and DeepSeek update their beliefs following major news events, earnings announcements, or policy shocks.

b. U.S.-based ChatGPT is systematically more optimistic than China-based DeepSeek about Chinese firms. What about other US-based LLMs relative to other China-based LLMs?

Gemini Claude Qwen3 豆包大模型

Thank you