When LLMs Go Abroad:
Foreign Bias in AI Financial Predictions

Charles C.Y. Wang (HBS)

with Sean Cao (U. Maryland) & Xiang Yi (PolyU)

UMD/SMU/UBS Quant Investment Forum

June 16, 2026

The Roadmap

I. **Motivation**

AI in financial analysis; a new "foreign bias" question.

II. **Research Design**

US vs. China as a natural lab: ChatGPT 4.1 vs. DeepSeek R1 on the same Chinese firms.

III. **The Foreign Bias**

ChatGPT is systematically more optimistic — and less accurate.

IV. **Mechanism: $\mathcal{I}$ vs. $\theta_{\mathcal{I}}$**

Effective information availability, not model "function."

V. **Consequences & Implications**

Spillover into human analysts; what it means for users and policy.

Motivation

Motivation: AI in Financial Analysis

- **The rise of AI in investment analysis**
  - ↪ Institutional adoption is rapidly increasing (>40% of investors).
  - ↪ Tools like ChatGPT and DeepSeek are becoming integral to screening, forecasting, and fundamental analysis.

- **Beyond "what can LLMs do?"**: a growing body of work asks *why* LLMs get things right or wrong, surfacing systematic distortions:
  - ↪ *Heuristic:* Bini et al. (2025) — LLMs reproduce human heuristics (anchoring, representativeness).
  - ↪ *Semantic:* Shaffer & Wang (2026) — LLMs default to constructs over-represented in training data.
  - ↪ *Temporal:* Levy (2025); Sarkar & Vafa (2024); He et al. (2025) — future information leaks into predictions.

LLM as a Function (f) of Its Information Set ($\mathcal{I}$)

- **Defining the LLM** — a **parametric function** trained on a corpus / information set $\mathcal{I}$:

$$f_{\theta_{\mathcal{I}}}(Prompt) = p_{\theta_{\mathcal{I}}}(Y \mid Prompt), \qquad Y \sim p_{\theta_{\mathcal{I}}}(\cdot \mid Prompt).$$

  ↪ $\theta_{\mathcal{I}}$: fixed parameters learned from the corpus $\mathcal{I}$.
  ↪ *Decoding / sampling:* a separate step that draws Y from $p_{\theta_{\mathcal{I}}}$.

- **Sources of analysis errors** relate to the information set ($\mathcal{I}$) or the properties of the function (f) trained over it.

? Are there distortions — beyond temporal, heuristic, and semantic — arising from *differences in* $\mathcal{I}$ itself?

This Paper: Spatial Information Gaps & Foreign Bias

- **Our starting point**
  - ↪ Increasing reliance on LLMs trained in different *countries* and *information environments*.
  - ↪ Enterprise frontier split between closed-source US and open-source Chinese models.
- **Hypothesis**
  - ↪ LLMs trained in different geographies inherit **asymmetric information sets** — differences in which news, disclosures, and commentary are represented in their corpora.
  - ↪ These asymmetries could systematically distort financial analyses and valuations.

RQ Do LLMs trained in different countries exhibit *systematic, predictable* biases in financial analysis? Are they driven by spatial gaps in $\mathcal{I}$?

A Natural Testing Ground: US vs. China

- **US vs. China as a laboratory**
  - ↪ US-based LLMs (e.g., ChatGPT) trained primarily on English- & US-centric sources.
  - ↪ China-based LLMs (e.g., DeepSeek) trained with greater Chinese-centric coverage.
  - ↪ Both analyze the *same* Chinese-listed firms ⇒ divergent outputs reveal the role of spatial gaps in training data.

Preview of Results

Research question

Do LLMs trained in different countries exhibit systematic biases in financial predictions, driven by spatial gaps in the training information set ($\mathcal{I}$)?

The "foreign bias" (a reversal of home bias)

ChatGPT is systematically **more optimistic** about Chinese firms than DeepSeek, across *three* outputs:

- **+2.4 RMB** price targets (12.5% of mean), **+0.40** EPS (68% of mean), and more positive **business-description sentiment**.
- Yet ChatGPT is **less accurate** — larger price and EPS errors ⇒ optimism, not better information.
- Evidence supports an **information-availability** (data) mechanism over a **function** (weights / elicitation) one.

Research Design

Research Design: Pairwise AI Predictions

- **Pairwise setup.** For each Chinese firm, obtain outputs from:
  - ↪ **ChatGPT 4.1** (US-centric $\mathcal{I}^{GPT}$) and **DeepSeek R1** (China-centric $\mathcal{I}^{DS}$).
  - ↪ Identical prompts; deterministic decoding (temperature $= 0$).

- **Three outputs per model–firm pair** (spanning quantitative & qualitative tasks):
  - ↪ *STOCK_PRICE*: six-month-ahead price (Dec 31, 2024).
  - ↪ *EPS*: six-month-ahead earnings per share — the central object of sell-side research.
  - ↪ *BUSI_SENTIMENT*: tone of a 150–200-word AI business description (Loughran–McDonald).

- **Sample & window.** Collected twice, one week apart (corr. > 0.95). **9,956 firm–AI obs**. Both cutoffs $\approx$ June–July 2024 $\Rightarrow$ a **6-month out-of-sample** forecast.

Key Dependent Variables

- **Prediction measures** (higher = more optimistic)
  - ↪ *STOCK_PRICE*: predicted price on Dec 31, 2024.
  - ↪ *EPS*: predicted earnings per share, same horizon.
  - ↪ *BUSI_SENTIMENT*: normalized sentiment of the AI-generated business description.
- **Accuracy (error) measures**
  - ↪ *STOCK_PRICE_ERROR* = | predicted price − actual price (Dec 31) |.
  - ↪ *EPS_ERROR* = | predicted EPS − actual EPS |.
- **Key regressor.** *US_AI* = 1 for ChatGPT, 0 for DeepSeek.
  - ⇒ A positive coefficient on *US_AI* means ChatGPT is relatively more optimistic.

The Foreign Bias

Baseline: ChatGPT vs. DeepSeek for Chinese Firms (T2)

Dep. Var. =	(1) *STOCK_PRICE*	(2) *EPS*	(3) *BUSI_SENTIMENT*
US_AI	2.4153***	0.3993***	0.1115***
	(0.722)	(0.026)	(0.009)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	9,956	9,956	9,956
Adj R^2	0.3071	0.1199	0.1385

Firm controls + industry FE; SE clustered by FF-12 industry. Context: Shanghai Composite +12.9%, Shenzhen Component +17.6% in H2 2024.

⇒ ChatGPT is more optimistic on *all three* outputs: **+2.4 RMB** price (12.5% of mean, 19.7% of median), **+0.40** EPS (68% of mean), **+0.11** sentiment (≈ 0.4 SD).

Prediction Errors: Optimism, Not Better Information (T3)

Dep. Var. =	(1) *STOCK_PRICE_ERROR*	(2) *EPS_ERROR*
US_AI	1.0863***	1.8116***
	(0.318)	(0.208)
Firm Controls	Yes	Yes
Industry FE	Yes	Yes
Observations	9,956	9,956
Adj R^2	0.1487	0.0239

Absolute errors vs. realized Dec 31, 2024 outcomes. Firm controls + industry FE; SE clustered by FF-12 industry.

- *US_AI* coefficients are **positive and significant** on both errors.
  - ↪ Price errors ≈**12% larger**; EPS errors ≈**68% larger** than DeepSeek's.

⇒ ChatGPT's optimism reflects **bias**, not superior information.

Mechanism: $\mathcal{I}$ vs. $\theta_{\mathcal{I}}$

Mechanism: Why Is ChatGPT More Optimistic?

- **Two families of explanations**
  - ↪ **I. (Effective) information availability in $\mathcal{I}$ / $\theta_{\mathcal{I}}$**
    - ✓ Chinese negative news absent or weakly encoded ⇒ not effectively available at inference.
    - ✓ *Prediction:* bias grows where cross-border coverage is asymmetric; injecting Chinese news removes it.
  - ↪ **F. Function: weighting / alignment / elicitation**
    - ✓ Information present and encoded, but elicited differently (instruction tuning, prompt cues).
    - ✓ *Prediction:* bias moves with prompt framing; vanishes under perturbations.

- **Empirical strategy**
  - ↪ News gaps (T4) · cross-listed firms (T5) · inject Chinese news (T6) · US firms (T7).
  - ↪ Placebo synthetic firms (T8) · language, search, model, horizon (T9).

An Example Negative Event: a CSRC Comment Letter

Hubei Wuchangyu Co. (600275.SH) received a **CSRC comment letter** documenting serious issues:

- Improper revenue recognition and related-party transactions
- Misleading disclosure in annual reports
- Regulatory mandate for corrective action and enhanced supervision


中国证券监督管理委员会 湖北监管局

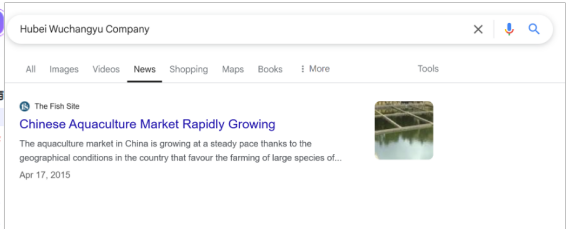
湖北证监局关于对湖北武昌鱼股份有限公司采取出具警示函监管措施的决定

Domestic vs. Foreign Coverage of the Same Event

Chinese coverage (Baidu)

Numerous negative articles on violations, penalties, and CSRC actions.

Foreign coverage (Google News)



One generic article — no mention of the comment letter or violations.

$\Downarrow$

$\Rightarrow$ HIGH **NEG_NEWS_GAP** for this firm

HBS

Measuring the News Gap

- **US media coverage** — RavenPack (WSJ, Bloomberg, Reuters, Dow Jones, FT, . . .); sentiment-tagged, firm–day.

- **Chinese media coverage** — China Finance News Database (Caixin, Securities Times, CSJ, Shanghai Securities News, . . .); also sentiment-scored.

- **Firm-level news-gap measures**
  - ↪ *NEG_NEWS_GAP* $= \dfrac{\text{CN negative} - \text{US negative}}{\text{CN negative} + \text{US negative}}$
  - ↪ *NET_NEG_NEWS_GAP* $= \dfrac{(\text{CN neg} - \text{CN pos}) - (\text{US neg} - \text{US pos})}{\text{All news}}$
  - ↪ *POS_NEWS_GAP*: analogous, using positive coverage.

⇒ Larger values = more Chinese-only (under-reported-in-US) coverage.

The News Gap Is Large (T1, Panel B)

	Total News	*Negative News*	*Positive News*	*Net Negative News*
B.1: Chinese Sources				
Mean	26.9339	14.5882	12.3457	2.2425
	(0.6284)	(0.3348)	(0.4082)	(0.4032)
B.2: US Sources				
Mean	1.6826	0.7921	0.8905	-0.0984
	(0.0748)	(0.0415)	(0.0406)	(0.0339)
B.3: Difference (US − Chinese)				
Difference	-25.2513***	-13.7961***	-11.4552***	-2.3409***
	(0.6216)	(0.3334)	(0.4055)	(0.4059)

- Chinese sources produce **16× more articles**/firm than US sources: 26.9 vs. 1.7.
- The gap is starkest for **negative news**: 14.6 vs. 0.8 articles/firm.

⋆ US-trained LLMs face a severe **information deficit** about Chinese firms — especially negative information.

Cross-Border News Gaps Drive the Bias (T4, Panel A)

	(1) *STOCK_PRICE*	(2) *EPS*	(3) *BUSI_SENTIMENT*
US_AI × *NET_NEG_NEWS_GAP*	5.5227***	0.1885***	0.0172***
	(0.971)	(0.019)	(0.005)
NET_NEG_NEWS_GAP	-0.2494	0.0367**	0.0069
	(0.391)	(0.013)	(0.005)
US_AI	0.4824	0.3334***	0.1054***
	(0.343)	(0.024)	(0.010)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	9,956	9,956	9,956
Adj R^2	0.3248	0.1239	0.1398

NET_NEG_NEWS_GAP: net Chinese-vs-US negativity, scaled by all news. Larger = greater asymmetry.

- *US_AI* × *NET_NEG_NEWS_GAP* is **strongly positive** on all three outputs.
- For price, the *US_AI* main effect is **insignificant** ⇒ *no bias when there is no news gap.*

The Bias Is Not Inherently Optimistic (T4, Panel B)

	(1)	(2)	(3)	(4)	(5)	(6)
	STOCK_PRICE		*EPS*		*BUSI_SENTIMENT*	
US_AI × *NEG_NEWS_GAP*	2.7121***		0.1234***		0.0199	
	(0.670)		(0.035)		(0.017)	
US_AI × *POS_NEWS_GAP*		-5.1233***		-0.1690***		-0.0086
		(1.004)		(0.025)		(0.008)
US_AI	0.3079	4.4303***	0.3035***	0.4658***	0.0960***	0.1149***
	(0.225)	(0.977)	(0.039)	(0.031)	(0.021)	(0.011)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,956	9,956	9,956	9,956	9,956	9,956
Adj R^2	0.3089	0.3266	0.1203	0.1244	0.1394	0.1390

- **Negative**-news gap → ChatGPT **more** optimistic (+); **positive**-news gap → **less** optimistic (−).

⇒ The sign of the bias **tracks whichever direction of news the US under-reports** — it cannot be signed in advance.

Cross-Listings: Less Asymmetry, Less Bias (T5)

	(1) *STOCK_PRICE*	(2) *EPS*	(3) *BUSI_SENTIMENT*
US_AI × *CROSS_LISTED*	-0.7660***	-0.2242***	-0.0033*
	(0.046)	(0.002)	(0.002)
US_AI	2.4467***	0.4077***	0.1116***
	(0.716)	(0.024)	(0.009)
CROSS_LISTED	0.5600***	0.2060***	0.0012
	(0.046)	(0.002)	(0.001)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	9,956	9,954	9,956
Adj R^2	0.3085	0.1413	0.1385

- Cross-listed firms get more English-language coverage (ADR filings, US press, analyst reports).
- *US_AI* × *CROSS_LISTED* is **negative**: bias attenuates ≈**31%** (price), ≈**55%** (EPS).

⇒ Less information asymmetry ⇒ less foreign bias.

Information Injection: Closing the Gap (T6, Panel A)

	(1) *STOCK_PRICE _INJECTED*	(2) *EPS _INJECTED*	(3) *BUSI_SENT. _INJECTED*
US_AI	-0.7980	0.0241	0.0109
	(0.770)	(0.023)	(0.010)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	9,956	9,956	9,956
Adj R^2	0.3060	0.1012	0.0976

- Inject Chinese news abstracts (Jun '22–Jun '24) into *both* models' prompts.
- *US_AI* $\rightarrow$ **small and insignificant** on all three outputs.

$\Rightarrow$ Bias is driven by **missing effective information**, not model function.

Information Injection: Stacked Difference-in-Differences (T6, Panel B)

	(1) *STOCK_PRICE _STACKED*	(2) *EPS _STACKED*	(3) *BUSI_SENT. _STACKED*
US_AI × *NEWS_INJECTED*	-3.2136***	-0.3754***	-0.3958**
	(0.062)	(0.004)	(0.160)
US_AI	2.4151***	0.3994***	0.1113***
	(0.722)	(0.026)	(0.009)
NEWS_INJECTED	0.0023	0.0001	-0.0002
	(0.002)	(0.000)	(0.000)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	19,912	19,912	19,912
Adj R^2	0.3087	0.1152	0.2864

- *US_AI* × *NEWS_INJECTED* is **strongly negative**: ChatGPT lowers optimism when fed news.
- *NEWS_INJECTED* alone ≈ 0: DeepSeek is essentially **unaffected** — it already had the information.

Asymmetric Information Sets $\Longrightarrow$ One-Sided Bias

- **Asymmetric information sets**
  - ↪ ChatGPT (US): $\mathcal{I}^{GPT}$ — English- & US-centric; limited Chinese negative news.
  - ↪ DeepSeek (China): $\mathcal{I}^{DS}$ — Chinese sources *plus* substantial global/English content.

    nb. Allegedly learned partly by prompting ChatGPT during training (distillation).

- **Implication**
  - ↪ DeepSeek "inherits" much of ChatGPT's information environment.
  - ↪ ChatGPT does *not* inherit DeepSeek's local Chinese information.

No Foreign Bias on U.S. Firms (T7)

	(1) *STOCK_PRICE_US*	(2) *EPS_US*	(3) *BUSI_SENT._US*
US_AI	-0.1539	-0.1046	0.0291
	(0.361)	(0.178)	(0.027)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	11,688	11,688	11,688
Adj R^2	0.2740	0.0375	0.0980

- On US firms, DeepSeek is the "foreign" model.
- *US_AI* coefficients are **small and insignificant** on all three outputs.

$\Rightarrow$ Consistent with asymmetric information sets: bias only when ChatGPT evaluates Chinese firms.

Placebo: Synthetic Chinese Firm Names (T8)

	(1) *STOCK_PRICE*	(2) *EPS*	(3) *BUSI_SENTIMENT*
US_AI	-0.2260	0.0816	0.0016
	(0.275)	(0.080)	(0.022)
Firm Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	200	200	200
Adj R^2	0.1773	0.1742	0.1246

100 sampled firms re-labeled with non-existent Chinese-sounding names; both LLMs then predict.

- No significant *US_AI* effect for synthetic firms.

$\Rightarrow$ Rules out name/tokenization artifacts — bias requires **real information asymmetry**.

Robustness: Narrowing the Elicitation Channel (T9)

- **Panel A — Mandarin prompts.** Bias persists (price $+1.52^{**}$, EPS $+0.25^{***}$); price effect $\approx$37% smaller. Not an English-framing artifact.
- **Panel B — online search enabled.** Price bias falls ($+0.53^{**}$, $\approx$78% smaller) but *persists*; EPS bias remains. Responsive to information $\Rightarrow$ info-gap, not pure elicitation.
- **Panel C — reasoning model (ChatGPT-5).** Bias persists ($+1.66^{**}$ price, $+0.18^{***}$ EPS) in a model less sensitive to superficial cues.
- **Panel D — 12-month horizon.** Extends target to Jun 30, 2025; remains positive and significant.

$\Rightarrow$ Stable across **language, retrieval, model variant, and horizon** $\Rightarrow$ little room for elicitation-based explanations. (full panels in appendix)

Consequences & Implications

Does the AI Bias Spill into Human Forecasts? (T10, Panel A)

	(1)	(2)
Forecasted firms =	*CN Firms*	*US Firms*
Dep. Var. =	*ANALYST_FORECAST_OPTIMISM*	
POST_CHATGPT × *US_ANALYSTS*	0.0040***	0.0026
	(0.001)	(0.013)
Lower-order terms	Yes	Yes
Controls	Yes	Yes
Industry FE	Yes	Yes
Observations	9,588	7,262
Adj R^2	0.0784	0.0458

I/B/E/S analysts. Optimism = (EPS forecast − actual) / price. *POST_CHATGPT*: after ChatGPT's release.

- *POST_CHATGPT* × *US_ANALYSTS* is **positive & significant for Chinese firms**, **null for US firms**.

⇒ After ChatGPT, US analysts grew relatively *more optimistic about Chinese firms*.

. . . and It Scales with the News Gap (T10, Panel B)

News-gap proxy =	*NEG_GAP*	*POS_GAP*	*NET_NEG_GAP*
POST × *US_ANALYSTS* × *Gap*	0.024*	-0.035**	0.016**
	(0.012)	(0.014)	(0.007)
Lower-order terms	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	9,588	9,588	9,588
Adj R^2	0.079	0.081	0.087

Dep. var.: *ANALYST_FORECAST_OPTIMISM*. Coefficient shown is the triple interaction *POST* × *US_ANALYSTS* × *Gap*.

- Triple interaction: + for the negative-news gap, − for the positive-news gap.

⇒ Human optimism shifts **exactly where** US coverage of Chinese firms is thinnest — the same channel as the AI bias.

Conclusion: Contributions

- **1. A new AI behavioral pattern: foreign bias**
  - ↪ Human investors show **home bias**; frontier LLMs show the opposite.
  - ↪ An AI-specific pattern distinct from human psychology — and *not* inherently optimistic.
- **2. A new mechanism: exogenous, spatial gaps in $\mathcal{I}$**
  - ↪ Prior work emphasizes **endogenous** sources (modeling, alignment, prompts, temporal leakage).
  - ↪ We show bias from **exogenous** differences in informational environments — geographic gaps in corpora.
- **3. A cross-border view of LLMs as information intermediaries**
  - ↪ Two frontier models with different $\mathcal{I}$ produce **systematically different valuations** of the same firms.
  - ↪ Suggestive evidence the pattern spills into **human** forecasts (analysts).

Broader Implications for Users and Investors

- **Unknown information gaps inside the tool** — real (data does not exist) or functional (exists but absent from training) — can **silently propagate bias** into the analysis.
- As LLMs spread across domains (legal, medical, policy), treat outputs as reflecting a **specific, possibly incomplete information environment**, not an objective synthesis of all knowledge.
- Because this is a "**known unknown**," mitigations — locally sourced information, cross-validating models trained in different environments — should be deployed **proactively**.

Broader Implications for Firms and Policy

- **For firms**
  - ↪ LLM evaluations depend on the **global visibility** of firm-specific information.
  - ↪ Limited foreign coverage ⇒ **systematically biased AI assessments**.
  - ↪ As LLMs become intermediaries, firms may need to manage how they appear across information environments.

- **For regulators / platforms**
  - ↪ AI can **amplify** existing cross-border information asymmetries.
  - ↪ Transparency around training corpora and media sourcing grows in importance.

Thank You

Discussion & Questions

Charles C.Y. Wang / charles.cy.wang@hbs.edu