Financial Stability Implications of Generative AI: Taming the Animal Spirits

by Anne Lundgaard Hansen and Seung Jung Lee
Discussion by Haoyang Liu

CEMLA/Dallas 4th Fed Fin Stability Conf

The views expressed are mine and do not necessarily reflect those of the Federal Reserve Bank of Dallas, or the Federal Reserve System.

Summary of Paper

- Increasing usage of Generative AI in asset management (Sheng et al. RFS forthcoming).

- Widely shared concern:
  - Gen AI $\Rightarrow$ Herding $\Rightarrow$ Financial Instability

- The authors carefully layout the hypotheses:
  - Ambiguous signs for (Gen AI $\Rightarrow$ Herding) and ($\Rightarrow$ Instability), e.g.
    - AI are computer programs and might display less human bias.
    - AI might reflect human bias because they are trained on human data (Stochastic parrots)

- Present experimental evidence: in general, AI agents less likely to herd than human.

General Feedback

- Important paper. Concern about herding driven by AI is real.

- Very thorough empirical analysis and careful discussion.

- Challenge (for all experimental studies): Map experiments to the real world.

- Facing gaps between experiments and the real world, researchers tend to argue for theoretical contributions, e.g. to psychology or neural science.

- Harder for this paper because it examines bias in machines instead of human.

- Pitch really matters:
  - Is financial stability the best pitch?
  - How about focus on human bias in Gen AI?

Map experiments to the real world

- AI versus human. How about AI + human?
  - JFE 2025 best paper: Cao et al. "From Man vs. Machine to Man + Machine: The Art and AI of Stock Analyses"

- In the real world: what would be considered private information for Gen AI agents?
  - Almost all information available to AI is public
  - What would be not considered as herding?

- Human benchmark is from pre-2009, has it changed?

Is financial stability the best or the only pitch?

- Studying financial stability does require realistic experiments.
    - Other behavioral finance experiments, e.g. on extrapolative beliefs, do not always need to be realistic.

- The authors acknowledge challenges in making financial stability statements:
    - How AI evolves over time.
    - AI adoption and how AI is used.
    - "In addition, the interaction between human and AI traders becomes crucial, as their combined behavior could either amplify or dampen market movements in unpredictable ways."

Findings about human bias in Gen AI is fascinating

- "They have inherited elements of human intuition and bias."

- Traditional algorithm trading would not show any human bias.
- Can the authors study: how to remove human bias in Gen AI models?

- AI models are evolving
- Can the authors do more with how Gen AI's human bias evolves?

✦ **AI Overview**

Yes, many users report that GPT-5 feels less "human" than GPT-4o, describing it as more formal, distant, and robotic. While GPT-5 has improvements in areas like reasoning and accuracy, the perceived loss of emotional warmth and personality has led to user disappointment and a sense that conversations are less natural and more rigid compared to GPT-4o's more "friendly" or "conversational" tone. 🔗

Is there a simple explanation for the results of less herding?

- Instead of AI reasoning through the logic and choose to not herd, could it be that AI is explicitly instructed to not herd?
- AI domain knowledge is trained by professionals in the field

OpenAI Looks to Replace the Drudgery of Junior Bankers' Workload

- Data providers for AI firms: ScaleAI, Surge, Mercor, Turing, etc.

Conclusion

- My takeaway: despite rising concerns that AI will lead to more herding, it is not true in experiments.

- Empirical analysis is very thorough.

- Positioning of the paper matters:
  - Financial stability?
  - Human bias in machines?