FINANCIAL INNOVATION

IMPLICATIONS FOR PAYMENTS AND POLICY

FEDERAL RESERVE BANK OF KANSAS CITY
JACKSON HOLE, WYOMING
AUG. 27-29, 2026

RAGHURAM RAJAN, PROFESSOR, UNIVERSITY OF CHICAGO

FEDERAL RESERVE BANK OF KANSAS CITY

Discussion of Brunnermeier's "Artificial Intelligence and the Brave New World in Finance"

Raghuram G. Rajan
University of Chicago Booth School

Outline of Discussion

1

Main points of paper.

2

My selective concerns about the analysis

3

Brunnermeier's policy suggestions.

Main points of paper

Finance exposed to AI revolution because of reliance on data, information [and incentives]

Three problems with AI

1

Non-alignment: Paper clip problem (Bostrom, Goethe, and Lucian of Samosata)

2

Non-explainability: Cannot tell why a loan application is rejected

3

AI Dominance: Most current AI models know more than any human agent and are more "intelligent" on specific issues

Asymmetric understanding

"One party anticipates the other's responses to relevant situation…more reliably than the other anticipates its responses"

Easy for AI to model humans but not vice versa

Humans at disadvantage

Human principal/central bank	*vs*	AI agent
Human trader	*vs*	AI trader

Our standard economic models don't handle such AI agents well

Example 1: Principal Agent relationship

- Traditionally, only don't know human agent's type/effort (hardworking/lazy)
- Traditionally, monitoring of the agent helps improve incentives

Example 2: Prices are informative

- Can invert from prices to information knowing how human traders maximize giving information
- Inversion hard if you don't know what AI traders maximize

Our standard economic models don't handle such AI agents well contd.

Example 3: Collusion in trading/Coordination

- AI agents learn to coordinate trading without instruction (Dou, Goldsteing, Ji (2024))
- Q learning algorithms learn trigger price strategies
- **Whom do you hold responsible?**

My assessment and (selective) concerns about the analysis

Assessment

1

Timely big picture paper by an academic leader in finance.

2

Suggests many fruitful areas for research and analysis

Concerns

1. **Is it humans vs AI or is it (humans plus AI) vs (humans plus AI)?**

2. **Are humans less susceptible to misalignment and lack of explainability than AI?**

3. **Brunnermeier worries about economic models becoming less useful with asymmetric understanding**

How we deal with these concerns today

Non-explainability:

Capital, collateral, margins

Non-alignment/Agency :

Skin in the game, supervision

AI Dominance:

Limits on scale and leverage, diversity of participants, competition

- Keep these in mind as we discuss policy prescriptions

Other possibilities not focus of paper

Can we get more alignment and more explainability from AI models?

- Constitutional AI
- Increase interpretability of AI models

Brunnermeier admits he focuses on the downside only

- Can AI keep its head when humans are panicking and provide liquidity?
- Isn’t the greatest danger of correlated AI behavior?

- **Portfolio insurance in 1987**

How likely is correlated behavior?

Prompt: Can you write this more elegantly?

Mary had a little lamb,
Its fleece was white as snow;
And everywhere that Mary went
The lamb was sure to go.

How likely is correlated behavior?

ChatGPT	*Mary had a little lamb,* *Whose fleece was white as winter snow;* *And everywhere that Mary wandered* *The faithful lamb was sure to go.*
Claude Sonnet 5	*Mary kept a lamb of tender grace,* *Whose fleece was pale as winter's morn;* *And wheresoever she set her pace,* *Faithfully the lamb was borne along.*
Gemini	*A gentle, snowy fleece adorned Mary's cherished lamb. Bound by a quiet devotion, it followed faithfully in her every footstep.*

Implication

Heterogeneity in AI architecture and training weakens coordination and predictability (Anand et al. (2026))

Bottom line: The real world, and humans, are plenty complicated even today. Will asymmetric understanding of AI make it qualitatively different?

Is the real concern whether AI will operate in new ways we have not contemplated yet?

Brunnermeier’s policy suggestions (intended to generate debate)

Suggestions for financial regulators

1

Segmented markets

2

Collusion regulation (if AI collusion becomes more of a problem)

3

Human authorization

Suggestions for central banks

Market participants know the reaction function of central bank but not vice versa:

1

Market participants set traps for central bank

2

Central bank should follow robust max-min approach

Central bank policy changes

1. **Rules vs discretion**
2. **Communication**
3. **Don't work through markets, work directly**
4. **Align AI through supervisory intervention**

Societal changes

Rightly worry about excessive AI concentration (power and risk)

Measures deal with ownership concentration but exacerbate logical concentration

- **Doesn't diversity across model structure reduce some of the other risks?**

Bottom Line

1

Lots to think about and prepare for.

2

Markus does a wonderful job of raising concerns.

3

But policy changes require more experience and debate.