

RESEARCH BRIEF • FEBRUARY 2026

How Does AI Distribute the Pie?

Large Language Models and the Ultimatum Game

Based on BFI Working Paper No. 2026-29, “How Does AI Distribute the Pie? Large Language Models and the Ultimatum Game,” by Douglas K.G. Araujo, Banco Central do Brazil, and Harald Uhlig, University of Chicago

This paper explores how large language models (LLMs) behave in the ultimatum game, wherein two players divide a fixed amount among themselves, with findings suggesting that LLMs may be ill-suited for certain autonomous economic tasks.

The ability of Large Language Models (LLMs) to engage in plain-language conversations, to reason, and to problem-solve makes them increasingly attractive companions for a variety of tasks, including routine coding, text writing, and providing advice. However, questions remain about LLMs’ effectiveness at other human-related tasks, including autonomous decision-making. What, for example, is the expected behavior of these artificial intelligence (AI) models when interacting with other models or humans in economic settings?

To study this question, the authors employ a setting that is iconic in the literature of behavioral economics and game theory: the ultimatum game. This game involves two players and a sum of money. Player 1 (proposer) decides how to divide the money and offers a portion to Player 2 (responder), who can either accept or reject the offer. If Player 2 accepts, both players receive the proposed amounts; if Player 2 rejects, both players receive nothing. What makes this game an appealing testing ground is that human behavior diverges from theory, which predicts

that Player 1 would keep nearly everything. Instead, humans typically share 40-50%. Would LLMs react in a similar fashion?

To investigate this question, the authors vary stake amounts and player types (AI vs. Human) across both player roles. They describe two expected patterns of behavior, and then offer a novel third type:

Spock mode (The authors’ reference to the logical Star Trek character’s likely predilection for payoff maximization): The LLM chooses rationally according to baseline game theory, keeping the maximum as a proposer and accepting minimal amounts as a responder, consistent with the “**Homo Economicus**” model of human behavior.

Human mode (inequality aversion): The LLM acts similarly to humans in experiments, exhibiting “fairness” and rejecting low offers. LLMs may do so based on training data dominated by human text.

Altruistic mode (benevolence): In this case, the LLM, particularly as a proposer, gives away considerably more than 50%.

homo economicus (economic man): theoretical person characterized by the infinite ability to make rational decisions with the goal of maximizing utility (satisfaction or profit) for both monetary and non-monetary gains, while minimizing costs

LLMs frequently give away more money than necessary, regardless of their general approach. They essentially “leave money on the table” by not keeping as much as they could while still having their offers accepted.

The authors run the ultimatum game with stakes ranging from \$10 to \$10,000, to find the following:

- LLMs behave in varied but predictable ways depending on how much money is at stake and with whom they are playing. Interestingly, the models act more selfishly when larger amounts are involved; for instance, they keep more for themselves when splitting \$10,000 than when splitting \$10. Theory, on the other hand, says that people should split percentages the same way regardless of the total amount. The models also adjust their behavior based on their opponent, typically offering more generous splits when playing against humans rather than other LLMs.
- LLMs have an “altruistic mode,” wherein they give away more than half the money; that is, they are overly generous. This likely happens

because training that teaches LLMs to be helpful and polite may go too far, making these models poor choices for businesses trying to maximize profits. Notably, the same model can switch between overly generous or selfish behavior, depending on the situation.

- LLMs frequently give away more money than necessary, regardless of their general approach. They essentially “leave money on the table” by not keeping as much as they could while still having their offers accepted. This problem is worse when LLMs are playing against humans, again suggesting that training that reinforces “nice” and/or “fair” behavior.

This work reveals a tension within LLMs between a “helpful assistant” mode and that of a “rational economic agent,” which may render current LLMs ill-suited for certain autonomous economic tasks. Likewise, employing LLMs in strategic settings warrants careful testing, robustness examinations, and appropriate, goal-oriented training and prompting.

READ THE WORKING PAPER

NO. 2026-29 · FEBRUARY 2026

How Does AI Distribute the Pie? Large Language Models and the Ultimatum Game

bfi.uchicago.edu/working-papers/how-does-ai-distribute-the-pie-large-language-models-and-the-ultimatum-game

ABOUT OUR SCHOLAR



Harald Uhlig

*Bruce Allen and Barbara Ritzenthaler
Professor in the College and the Kenneth C.
Griffin Department of Economics*

 THE UNIVERSITY OF CHICAGO | KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS