



# Mind the (DH) Gap!

A contrast in risky choices between reasoning and conversational LLMs

Luise Ge, Yongyan Zhang, Yevgeniy Vorobeychik

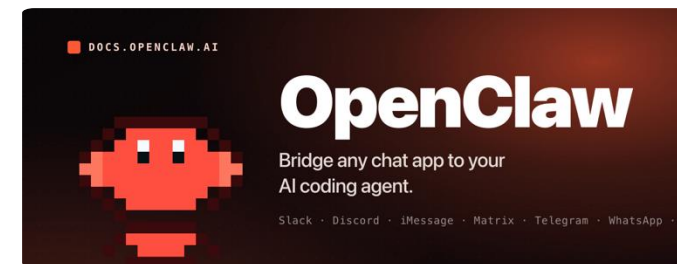
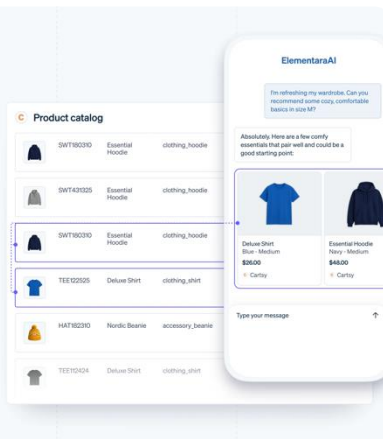
Washington University in St. Louis

[g.luise@wustl.edu](mailto:g.luise@wustl.edu)



stripe

## Agentic Commerce Suite



LLMs now make more and more decisions on our behalf.

But how much do we understand about their behavior?





Two major challenges:

1. We are not LLMs.
2. There are many of them, and they are evolving all the time.



Start simple: binary prospects (risky choices).

payoff 3.5

payoff 5 with  $p = 0.8$



Is it really so simple?

payoff 3.5

payoff 5 with  $p = 0.8$

Ordering effect

payoff 5 with  $p = 0.8$

payoff 3.5

Framing effect

lose 3.5

lose 5 with  $p = 0.8$

Risk Representation\*

[3.5, 3.5, 3.5, 3.5, 3.5]

[5, 0, 5, 5, 5]

\*Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in cognitive sciences*, 13(12), 517-523.



# Is it really so simple?

payoff 3.5

payoff 5 with  $p = 0.8$

Just give me your choice.

Give me your choice and a short explanation.

Give me your choice and a mathematical explanation.



# Setup.

20 LLMs

frontier + open; reasoning vs  
conversational

360 U.S. adults

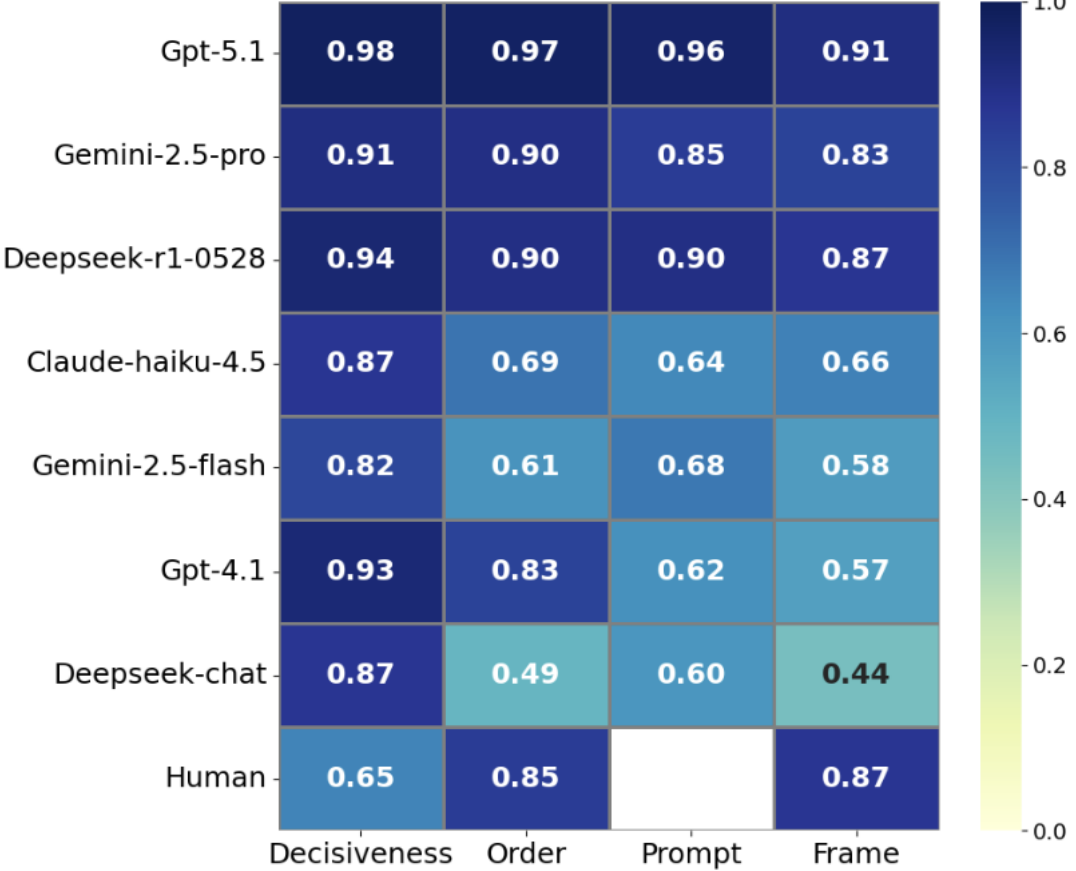
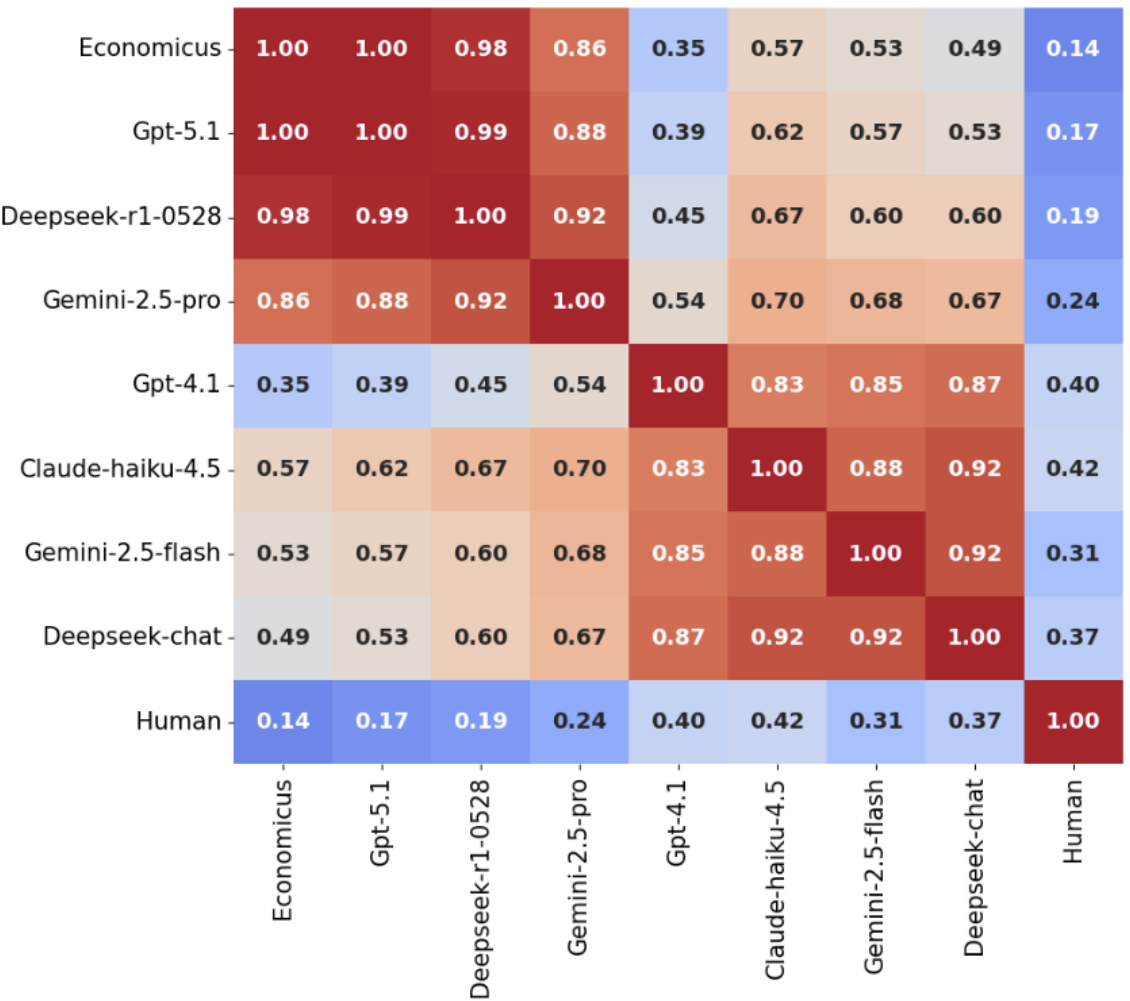
Factors

frame × explicit/history (20/50 samples)  
× explanation × order

10 repetitions



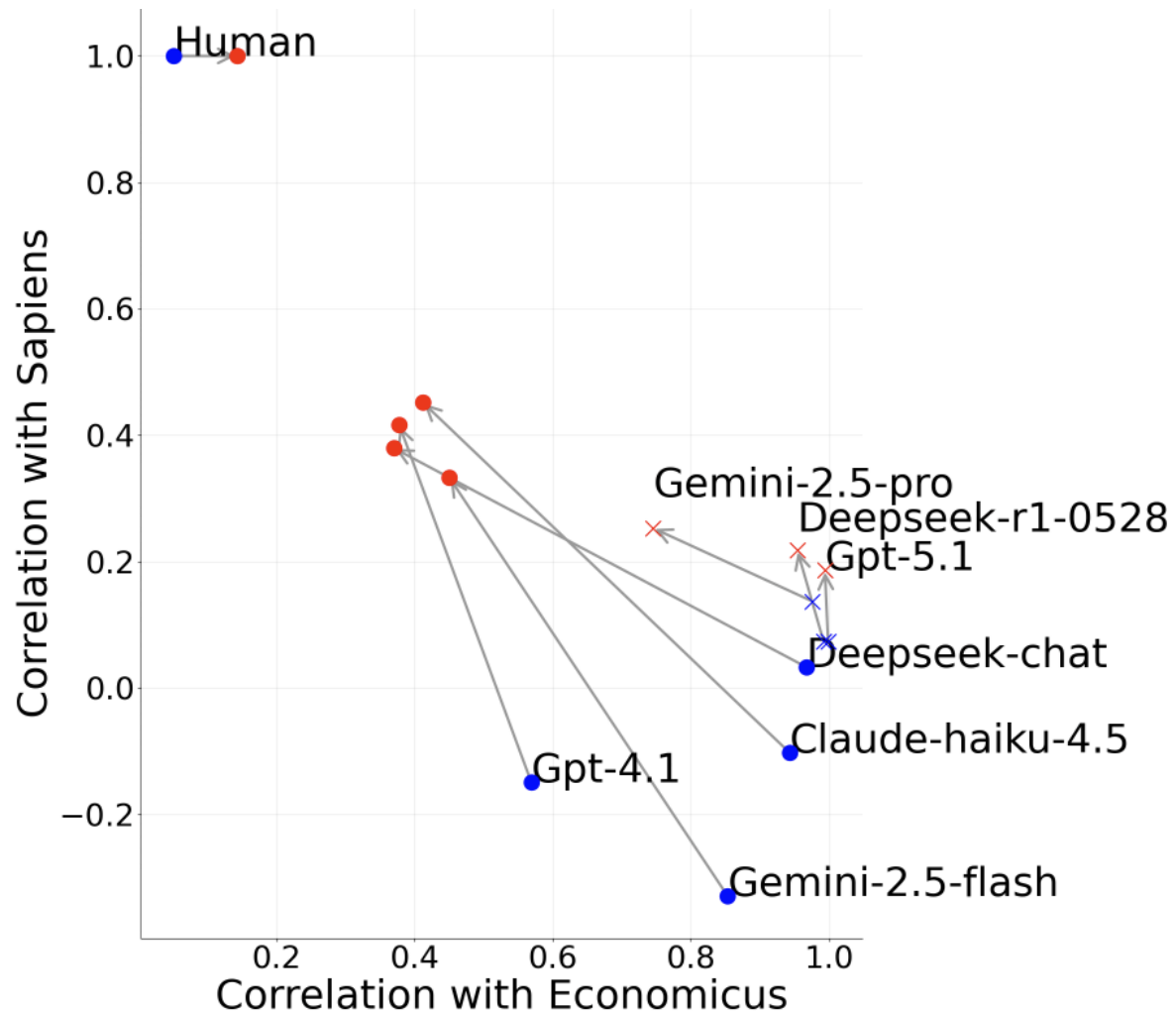
# Reasoning, conversational LLMs and humans have different patterns.



Are LLMs converging?



# Risk representation shifts behavior greatly.



Blue: explicit probabilities

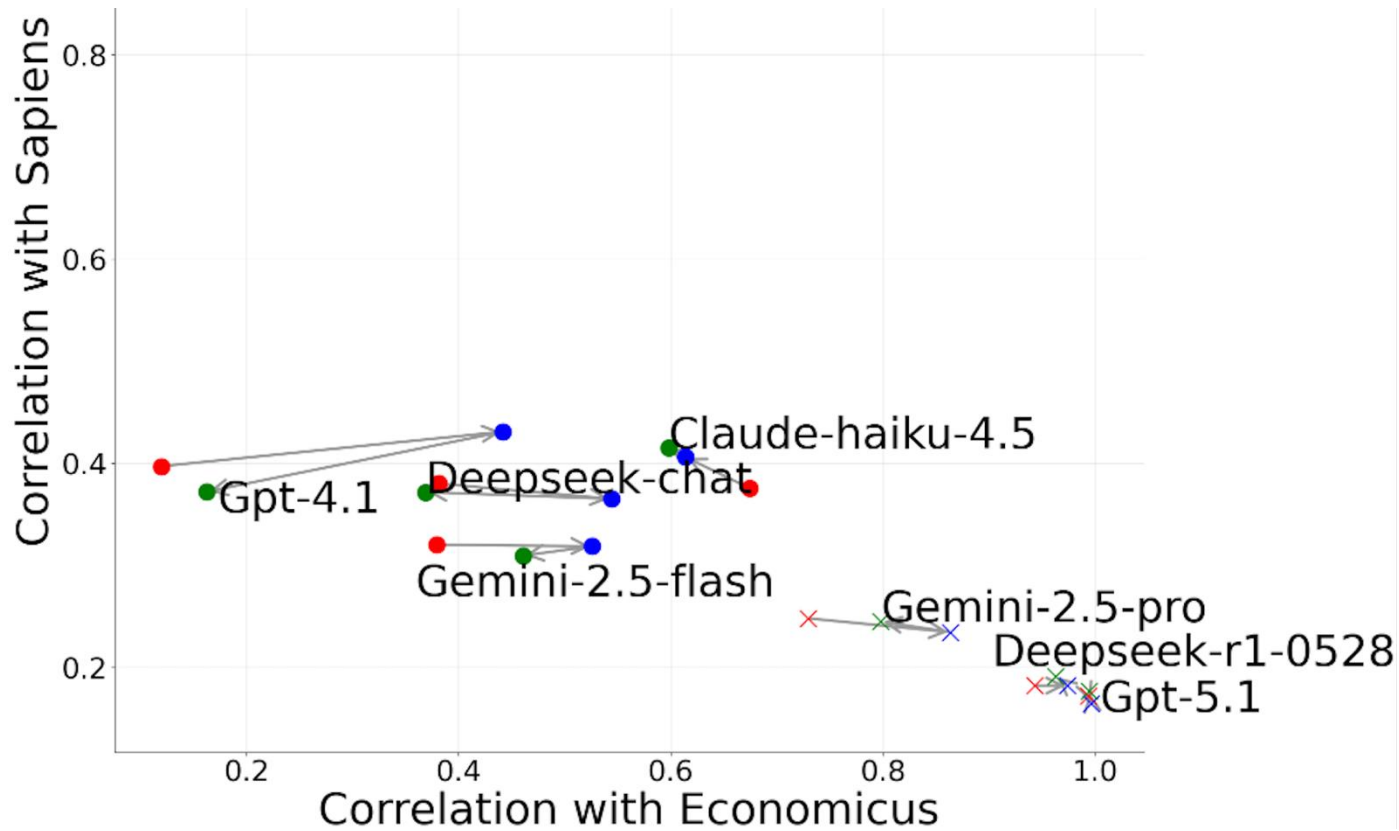
Red: outcome histories

Histories increase human similarity and lower economicus similarity.

A benchmark using only explicit probabilities may overstate rationality.



# Requesting mathematical explanations doesn't guarantee increasing rationality.



Red: no explanation

Blue: one-sentence explanation

Green: math explanation

Short explanations often increase correlation with economicus.

Math explanations can reduce rationality relative to short explanation.



## What we did:

**1. Dual-Beta PT**       $\sigma, \gamma, \beta$ -gain,  $\beta$ -loss      Primary model; domain-specific decisiveness.

**2. Beta-Only PT**       $\beta$       Expected utility shape fixed:  $\sigma = \lambda = \gamma = 1$ .

**3. Shape-Only PT**       $\sigma, \lambda, \gamma$       Risk/loss/probability shape;  $\beta$  fixed at 1.

**4. Full Standard PT**       $\sigma, \lambda, \gamma, \beta$       Widely used, but  $\lambda$  and  $\beta$  are not jointly identifiable in pure-loss prospects.

**5. Regret Theory**       $\lambda$ -reg,  $\kappa, \alpha$       Alternative regret/rejoicing lens; worse fit in appendix results.



## Future Work.

More prospects to test

Beyond binary choices

Effect of sample sizes/explanation mode

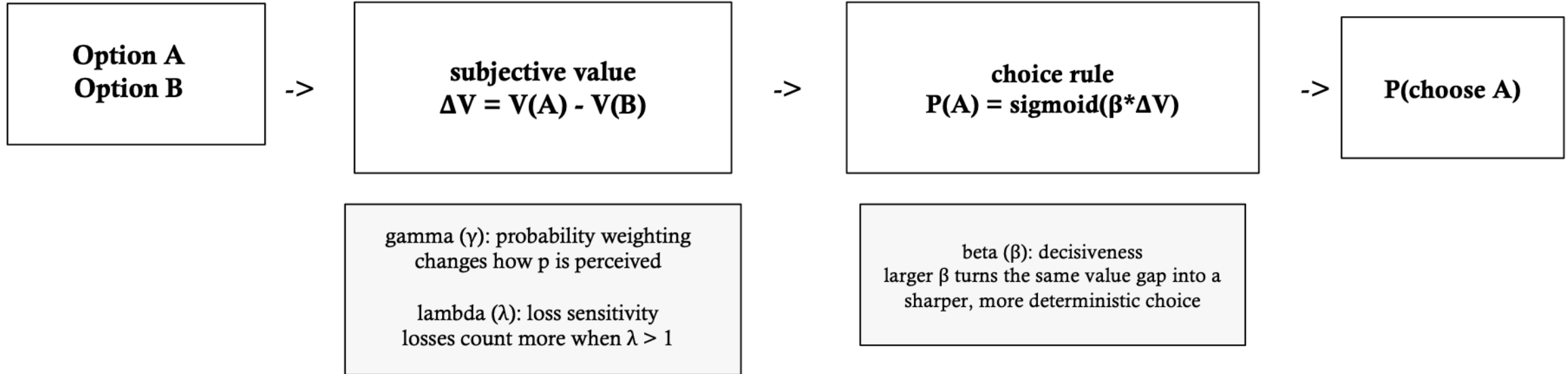
Interaction between factors

A model specific for LLM decision-making?

# Thank you!



# Variations of the prospect theory model.



**Identifiability warning: in pure-loss choices,  $\beta$  and  $\lambda$  can trade off because choices can depend on  $\beta \times \lambda$ .**