

Discussion of "How Rational Is AI Investment Advice? Risk-Return Relevance in AI-Based Investing"

Lin, Gilbert, Chu & Jacoby

DISCUSSANT: RUI DAI, UNIVERSITY OF SOUTH FLORIDA

2026 FMA ANNUAL MEETING | TAMPA, FL | OCTOBER 15, 2026
SESSION 70: AI AND FINANCIAL MARKETS

The paper in one slide

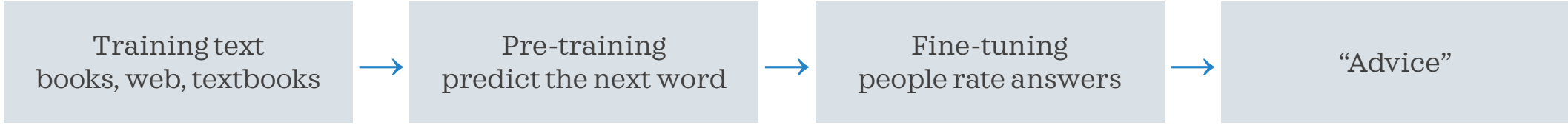
Question. Does AI advice follow textbook risk–return logic?

Design. 5,000 prompts with random market numbers and clients; GPT-5 splits $1,000 between a stock fund and a 2% bond.

Finding. More stocks when returns are high, fewer when volatility is high, but no response to the stock–GDP correlation ρ.

My take: timely question, clean design. The AI repeats its mean-variance training text.

Comment 1: An LLM mirrors its training text



The rationale is only as good as the input text. Its “reasoning” repeats what it read most.

Books are prized training data: firms scanned millions of print books, textbooks included; some used pirated libraries (2025 court records).

Finance text is mostly mean-variance: return versus risk. Recession hedging lives in specialist research; popular advice uses age rules.

So the pattern is predicted: return, volatility and age matter (Tables 3, 8, 9); ρ doesn’t. That reflects the corpus, not rationality.

Comment 1 (cont.): Test the training-text story

1. Measure the corpus. How often do mean-variance vs hedging ideas appear in textbooks, websites and SSRN abstracts? Do the AI's responses match?

2. Knows it, but doesn't use it? Ask GPT-5 why stocks that fall in recessions are riskier. If it knows but ignores it in advice, that's habit, not ignorance.

- Ask for a one-line reason with each allocation: does it mention the economy?

3. Prompt it explicitly: "use Merton's intertemporal CAPM", "treat stocks as a hedge".

- If it then responds to ρ, it knew all along, but only users who ask get the better advice.
- If not, it really can't.

4. Same text, different tuning. GPT-4.1 and GPT-5 read similar text, but only GPT-5 cuts stocks when volatility rises (Table 8): that's fine-tuning.

Without test 1, a training-data story can explain anything.

Comment 2: Compared with whom?

The issue. The paper grades GPT-5 against an ideal textbook investor, but most people and advisors also ignore how stocks move with the economy. Failing, or doing what everyone does?

Why it matters. "Irrational" only means something relative to the real alternatives: own judgment, a human advisor, or a robo-advisor.

- A self-driving car tested on an empty track tells you little about city traffic.

The textbook is less clear-cut: for the client in the prompt, theory suggests ρ may not change the allocation (appendix A1).

Suggestions

Compare with **people**: Chinco, Hartzmark & Sussman (2022) ran the same survey on humans.

Compare with **robo-advisors**: same client profiles, rule-based allocations.

Tell the AI the **client's job** ("construction worker") and see if the advice changes.

Report the **cost to the client in dollars**, not only p-values.

Comment 3: Why it matters, and for whom

From one client to the whole market

One client: a stock that falls when you lose your job is poor insurance.

Millions of clients: same model, same tilt. Mistakes add up instead of canceling out.

Price it first: what does following the AI cost a client per year, versus the textbook allocation and versus a human advisor or robo-advisor?

Gender: test it first

The paper never tests whether men and women differ.

Average advice is the same for both (Table 1C).

Test: same prompts, only the gender swapped.

Minor comments

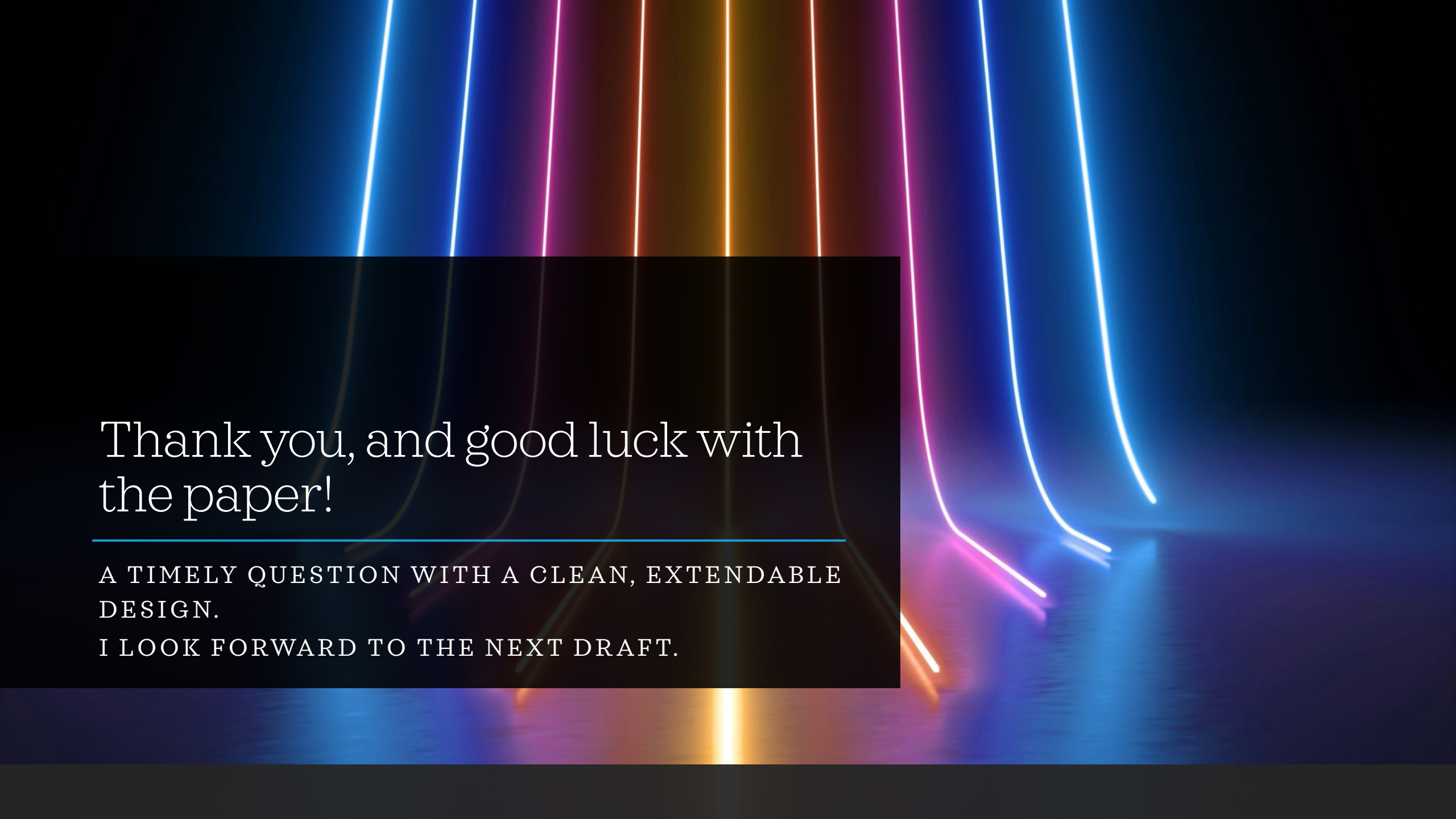
Re-check the dependent variable and standard errors. The "logit" share in Table 1 looks like the odds ratio (its maximum is 9,997; a log would stop at 9.2), and standard errors are clustered on only two groups (male / female). The main results depend on both.

Prompt (App. A.1): is the model advising a client or investing "for yourself"? The risk aversion used in Tables 4 and 9 is not in the sample prompt.

Let ρ go negative (footnote 3 rules it out): stocks that hedge the economy are the cleanest test.

Table notes say t-statistics are in parentheses; the numbers are p-values.

Title: "rational" → "consistent with textbook portfolio theory".



Thank you, and good luck with the paper!

A TIMELY QUESTION WITH A CLEAN, EXTENDABLE DESIGN.

I LOOK FORWARD TO THE NEXT DRAFT.

Appendix A1: What does theory predict for ρ?

For the prompt as posed, ρ does not enter. Only financial wealth, with expected return μ and volatility σ given: $w^* = (\mu - r) / \gamma\sigma^2$. In equilibrium μ already pays for covariance with consumption.

ρ matters only with other risk, such as a paycheck that moves with the economy (Bodie, Merton & Samuelson 1992; Viceira 2001). The prompt never mentions it. Under the corpus's assumptions, ignoring ρ is correct.

Fix: compute the textbook share for every prompt and compare it with the AI's; tell the model about the client's job risk and vary it; measure what a client loses by following the AI.

$$w^* = (\mu - r)/\gamma\sigma_s^2 \cdot (1 + H/W) - \rho \cdot (\sigma_y/\sigma_s) \cdot H/W$$

Change	Textbook (γ = 5)	GPT-5
+1 pp return	+17.3 pp	+7.7 pp
+1 pp volatility	−9.4 pp	−1.1 pp
+0.1 ρ	−1.2 pp × H/W	−0.9 pp

Change in % of $1,000 put in stocks. Textbook column: discussant's calculation at the median draws (appendix A2–A3); GPT-5: Table 2B, col. 5. σ_s, σ_y: stock, income (GDP) volatility; H, W: human capital, financial wealth.

Reading: GPT-5 moves the right way but too little, above all on volatility (a tenth of the textbook response at γ = 5, a quarter even at γ = 10). On ρ, theory allows anything from no response (no job income) to a sizable one; GPT-5's small cut sits in that range.

Appendix A2: the textbook benchmark

Setup (the prompt's investor): CRRA risk aversion γ; financial wealth W; a stock fund with expected return μ and volatility σ_s; a bond paying r = 2%.

Merton (1969, 1971): with financial wealth only, the optimal stock share is $w^* = (\mu - r) / \gamma\sigma_s^2$. ρ does not appear.

Adding labor income (Bodie, Merton & Samuelson 1992; Campbell & Viceira 2002, ch. 6): human capital H whose growth has volatility σ_y and correlation ρ with stocks holds an implicit stock position $\beta \cdot H$, with $\beta = \rho\sigma_y/\sigma_s$.

- Target stock exposure on total wealth: $(\mu - r)/\gamma\sigma_s^2 \cdot (W + H)$. Subtract the implicit $\beta \cdot H$ and divide by W:
- $w^* = (\mu - r)/\gamma\sigma_s^2 \cdot (1 + H/W) - \rho \cdot (\sigma_y/\sigma_s) \cdot H/W$
- Frictionless version (H tradable). Viceira (2001) solves the non-tradable case: more complex, same sign.

Sensitivities used in the table (rows 1–2 at H = 0, the prompt's investor; row 3 per unit of H/W):

- $\partial w^*/\partial\mu = 1 / \gamma\sigma_s^2 \quad \partial w^*/\partial\sigma_s = -2(\mu - r) / \gamma\sigma_s^3 \quad \partial w^*/\partial\rho = -(\sigma_y/\sigma_s) \cdot H/W$

Appendix A3: the numbers

Input	Value	Source
Stock volatility σ_s	10.75%	median draw, Table 1A
Excess return $\mu - r$	4.93% − 2% = 2.93%	median draw, Table 1A; bond in prompt
Income volatility σ_y	1.26%	median GDP volatility, Table 1A
Risk aversion γ	5	chosen: textbook share 0.0293 / (5 × 0.01156) = 51% ≈ GPT-5 mean 48%

Change	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$	GPT-5 (Table 2B, col. 5)
+1 pp return: $1/\gamma\sigma_s^2 \times 0.01$	+43 pp	+17.3 pp	+8.7 pp	7.74 × 0.01 = +7.7 pp
+1 pp volatility: $-2(\mu - r)/\gamma\sigma_s^3 \times 0.01$	−23.6 pp	−9.4 pp	−4.7 pp	−1.14 × 0.01 = −1.1 pp
+0.1 ρ: $-(\sigma_y/\sigma_s) \times 0.1$	−1.2 pp × H/W	−1.2 pp × H/W	−1.2 pp × H/W	−0.095 × 0.1 = −0.9 pp (p = 0.31)

Caveats: textbook numbers are local responses at the median prompt; GPT-5's are average slopes over all prompts, including 0% and 100% answers. σ_y uses GDP volatility as a proxy for income risk; H/W is unknown. The volatility gap survives any reasonable γ; the return gap does not (γ = 10 gives +8.7 vs +7.7).