Discussion of

Mimicking Finance

Lauren Cohen, Yiwen Lu,
Quoc Nguyen

Jonathan Reuter
Boston College & NBER

NBER Conference on Big Data, Artificial Intelligence, and Financial Economics

December 5, 2025

Very Big Picture

- **Motivational claim:** "Agents should not be compensated for aspects of their behavior and human capital that can be replicated in a low-cost way"
- Related questions of how AI/ML technology will impact
  - **Labor market** for asset managers (greater rewards for skill?)
  - **Product market** (more competition? less closet indexing?)
  - **Regulators/Intermediaries** (improved regulation or ratings?)
- Paper focuses on extent to which ML tools make stock-level trading by MFs **more predictable to econometricians**
  - Concludes by testing whether MF with less predictable trades outperform their peers... **as is mandatory for any MF study**

Me versus Machine

- Reading this paper immediately caused me to ponder my value-added as a discussant... and the optimal mixture of jokes versus insights
- So I read the paper, jotted down notes, and outlined a standard discussion
- Then I paid $49.99 and ran the paper through refine.ink
- I also asked ChatGPT to critique the paper, give me a list of missing citations, and rank it among Lauren's other papers
- AI tools responded to my prompts by focusing more on the trees whereas I tried to focus more on the forest

Empirical Strategy?

- Runs data on **5,434,702 fund-stock-quarters** through "Long Short Term Memory" (LSTM) model to predict buys, holds, and sales by **1,706 active equity mutual funds**, 1990-2023

  No, I'd never heard of **LSTM** models before either...

- Average prediction accuracy of **71%** with surprisingly small differences across styles and over time

- First set of tests seek to explain cross-sectional variability in fraction of correctly predicted manager/fund trades

- Remaining tests show that unpredictable trades predict outperformance relative to peers

My General Reaction?

- Provocative (early-stage) paper tackling fundamental question about extent AI will disrupt specific labor market
- Predicting general actions **(buy/hold/sell)** is a necessary... but not entirely satisfying first step... towards predicting specific actions **(how much and when)**
    - I can predict Lauren's presentation style and ability to publish well without being able to replicate either
- Want to focus more on how AI tools will change monitoring inside asset management firms
- Interpretating performance differences as arising from skill/effort requires more engagement w/ existing literature

More Specific Comments

1. Not sure how to think about model's predictions because I don't know how much **insiders** are learning from tools
2. How much do predictable differences in fund performance reflect differences in manager effort/skill?
3. How does any of this relate to manager compensation and risk of job loss?
4. While most cross-sectional differences are consistent with authors' predictions, my attempts to measure economic significance led to very small and VERY LARGE estimates

Bonus: Areas where AI outperformed me

Comment #1: Model?

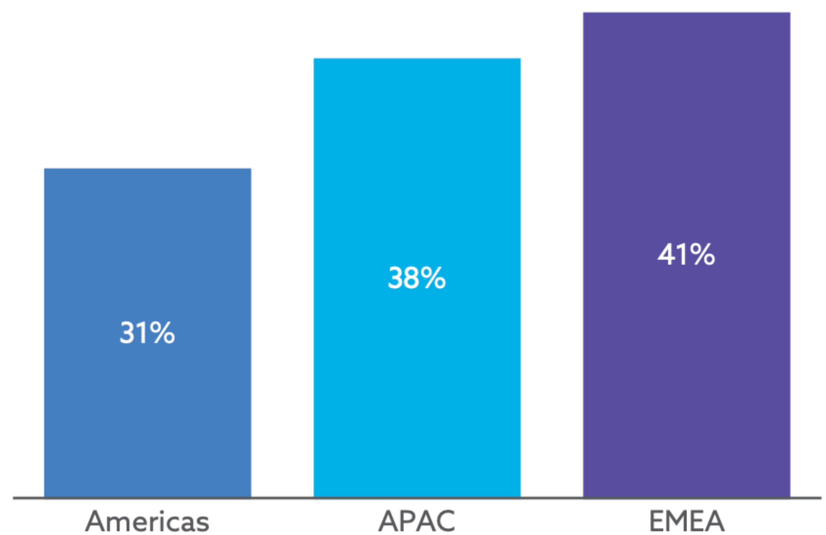
AI tools allow econometrician to identify predictable trades from public data → *industry moves from low-effort pooling equilbrium to high-effort separating equilibrium*

1. If we take model literally, and assume AI tools instantly increase monitoring inside asset management firms, post-AI period is 14 months between 11/22 and 12/23
2. Bigger issue? No distinction between inferences by
   A. Econometricians using public data
   B. Asset management firms using private data on why managers trade, how trades perform, alpha versus betas, impact of flows on returns, etc.

Comment #1 *(cont.)*

3. While AI tools are likely to increase monitoring inside asset management firms, they are more likely to narrow the gap between what academics (or competitors) can measure and what asset mgmt firms can already measure

Exhibit 4. Percentage of Respondents Who Have Used Generative AI Tools, Such as ChatGPT, in Their Daily Workflow by Region (*n* = 1,210)



Source: Brian Pisaneschi, "Unstructured Data and AI: Fine-Tuning LLMs to Enhance the Investment Process," CFA Institute (1 May 2024). https://rpc.cfainstitute.org/en/research/reports/2024/unstructured-data-and-ai.

Comment #1 *(cont.)*

4. Cross-sectional predictions regarding fund size, number of managers, manager tenure, and managerial ownership stake are plausible...

   ... but they ignore endogeneity of fund size and management team structure and extent to which these measures would change in post-AI equilibrium

Comment #2: Predictable Returns?

I'm not yet convinced that unpredictable trades are more likely to reflect high effort trades... and curious to learn more about how return differences relate to earlier studies

To what extent do differences in predictability reflect differences in portfolio characteristics or investor flows?

- Funds with more volatile investor flows may have more predictable trading and lower returns (Edelen JFE 1999)
- Tests for skill should focus on discretionary trades (Alexander, Cici, and Gibson RFS 2007) → are these more likely to be classified as unpredictable trades?

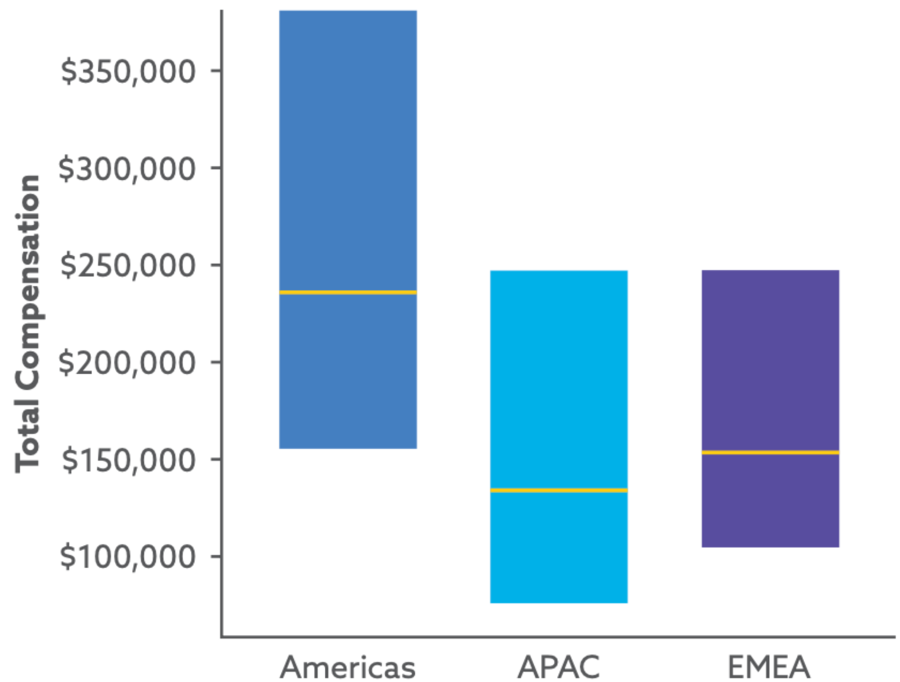
Comment #2 *(cont.)*

- Skill persists when funds hold relatively unique portfolios within their style (Hoberg, Kumar, Prabhala RFS 2017) → compare alphas rather than style-adjusted returns?
- Style drift should make trading less predictable relative to peers funds and drive up returns → compare alphas rather than style-adjusted returns?
- Two-way sort based on level of predictable trades and "return gap" which compares realized returns to those predicted in absence of trading (Kacperczyk, Sialm, Zheng RFS 2008)?

Comment #3: Compensation?

- Data on fund manager compensation is hard to come by
- 2024 CFA reports Q1-Q3 of total compensation in Americas, Asia-Pacific, and Europe, Middle East, and Africa
- Ideally, want to correlate compensation with paper's measure

Exhibit 23. Portfolio Management Total Compensation Quartiles by Region: All Seniorities (*n* = 4,557)



Note: The bars indicate high/low compensation quartiles, and the line in each bar denotes the median.

Comment #3 *(cont.)*

- Del Guercio, Genc, Reuter, Tran (2024) classify managers into three categories:

	Passive	Hybrid	Active
# Manager-years	531	411	10605
% Non-domestic equity	44.8%	44.5%	27.5%
# Funds (all types)	34.7	27.9	4.8
# Prospectus BMs	31.6	17.7	3.0
ETF manager	48.2%	22.6%	0.8%
Hedge fund manager	9.0%	26.3%	7.8%
MBA	43.5%	42.6%	60.6%
PhD	0.8%	4.9%	4.5%
SAT score (2003)	1192.9	1219.9	1273.8
Elite school	4.7%	14.3%	28.4%
Connected school	15.7%	23.3%	45.9%
Minority	20.7%	13.2%	8.9%
Female	16.0%	9.2%	10.4%
Quantitative major	19.9%	38.4%	20.0%

Comment #3 *(cont.)*

- Expect passive-only managers are most predictable (or easily replaced by algorithms) and active-only are least
- Income data that DGRT 2024 collected from mortgage applications are broadly consistent with this view

	Passive	Hybrid	Active
N	45	48	650
Mean income	\$272k	\$383k	\$751k
Median income	\$222k	\$254k	\$380k
Std dev income	\$202k	\$330k	\$1041k

Less predictable

- **Crucial next step would be to ask whether active-only managers with less predictable trades earn more**

Comment #4: Economic Significance?

- **Standard approach?** Multiply estimated coefficient by one-standard-deviation change in independent variable
- Authors report overall standard deviations when we want to be using (much smaller) within-fund standard deviations
  - *Manager Ownership:* *-0.0003 * 75.08 =* ***-0.0225***
  - *Number Managers:* *-0.0014 * 2.81 =* ***-0.0039***
  - *Total Net Asset:* *-2.1384 * 1391.91 = -2976.46*
  - *Fund Flow:* *-4.5290 * 108.74 = -492.48*
- No idea how to make sense of TNA and flow estimates
- Suggestion: estimate models that replace fund FEs with style-by-time FEs and be explicit about any normalizations

Refine.ink was all about the trees

- **22 comments** focusing on internal consistency
- Raises two issues with model (e.g., if routine tasks are free then additional assumptions are required to prevent $R = \infty$)
- Potential issues with proofs of Propositions 2 and 4
- Ambiguity in LSTM input tensor processing description "should be clarified for reproducibility"
- "precision and accuracy are distinct measures in multi-class classification" but being used interchangeably
- Authors reference N30-D filings but those were replaced by other filings between 2004 and 2023

ChatGPT was more balanced

Several comments overlap with mine; fourth seems legit

1. Questioned extent to which no trading is driving accuracy
2. Questioned whether predictable actions are low-skill
3. “The paper interprets poor predictability as ‘high skill,’ but reverse causality or omitted variables could drive the predictability–performance relationship. No instrument, decomposition, or structural estimation addresses this.”
4. “LSTMs designed for long-sequence temporal dynamics, yet the fund data are sparsely reported at quarter-ends and padded with synthetic balanced-panel structure” → value to additional model comparisons

ChatGPT was more balanced (2)

When I asked for missing citations, papers were reasonable but not always consistent with predictable trading = lower returns

- Lou (2012) for link between flows and future returns
- Kacperczyk, Sialm & Zheng (2005) because “predictability may correlate with specialization rather than skill”
- Also included Lauren’s 2005 “Judging Fund Managers by the Company They Keep” because it makes inferences about skill by comparing trading behavior of peer funds

Rank of this paper among all of Lauren’s papers? **9th**

Tony’s Unpredicted Contribution?