AlphaManager: A Data-Driven-Robust-Control Approach to Corporate Finance

Murillo Campello, Lin William Cong, Luofeng Zhou

Present by Long Zhen

Background

Causal Inference | Predictive Analysis + Prescriptive Analysis

- Econometrics
- → causal tree

- New X or New *f(·)*
- → supervised learning

- Operation management
- → reinforcement learning

- **How to apply AI techniques?**
- **How to combine the 3 paradigm?**

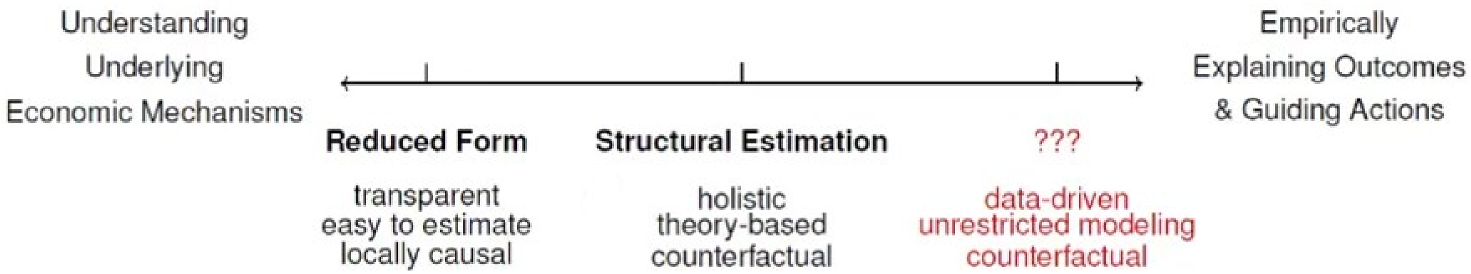
Background

- Modern AI: intelligence → solve problems and achieve goals in reality
- Three ways of AI:
  - Larger model, greater computation
  - More data, more computation; pre-training
  - Clever goal-oriented algorithm; knowledge, logic, and expertise

- AI in finance: handle big data, low signal-to-noise, nonlinearity…
  - Relying mostly on statistical properties of data, or off-the-shelf models
- What about corporate finance?

Motivation

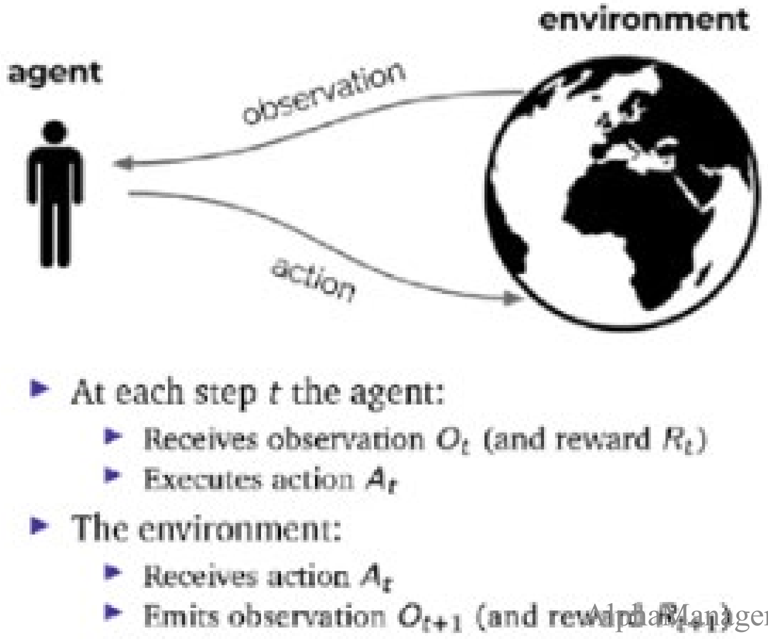
- What are managers' optimal decision-making?
  - Counterfactual statement:
    - Experiment & survey
    - Reduced-form models/ causal inference
    - Structural approach: need a theory for the general dynamics

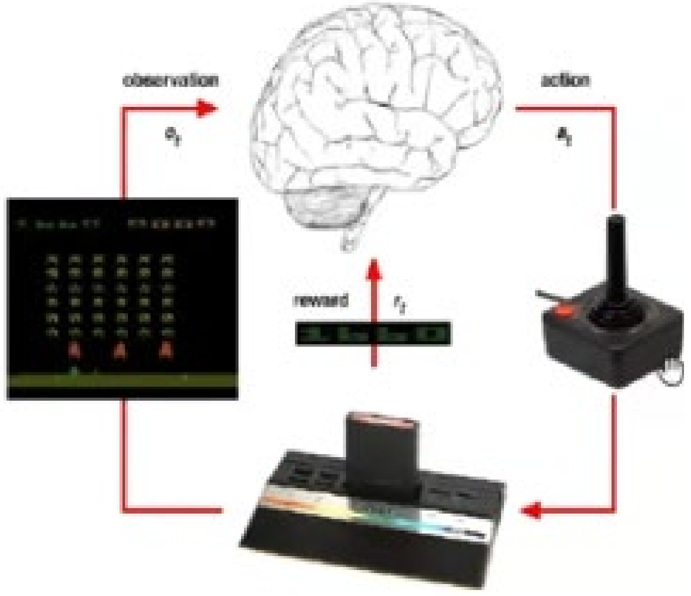


  - Manager actions interact with environment
    - High-dimension; interaction; fast evolving
  - → stochastic control problem:
    - Information set
    - Dynamically learned, high-dimensional, and nonlinear
    - Need robust counterfactuals

Reinforcement Learning

- Agenda**:** economically guided heuristic search and greedy search in a large decision/action space to achieve goals
- People learn by **interacting with the environment** to optimize rewards.





Research question

- How to develop a data-driven-robust-control system for corporate finance decision making?

- What they do:
  - Deep neural networks to accommodate high-dimension and nonlinearity
  - Offline deep reinforcement learning to incorporate managerial objective
  - Introduce ambiguity aversion

Contribution

- AI in finance
  - ML in CF
    - Textual analysis, e.g., Li et al., 2021. Cong, Liang, & Zhang, 2019, etc.;
    - Supervised learning, e.g., Erel et al., 2021, Lyonnet and Stern, 2022
  - New DDRC overcoming limitations and unifying framework.
- Robust Control:
  - Mostly theory, focus on macro time series rather than utilizing cross-sectional info (e.g., Ju and Miao, 2012).
  - Application in corporate finance.
- Artificial intelligence:
  - Goal-oriented search
  - Model-based offline RL (empirical).

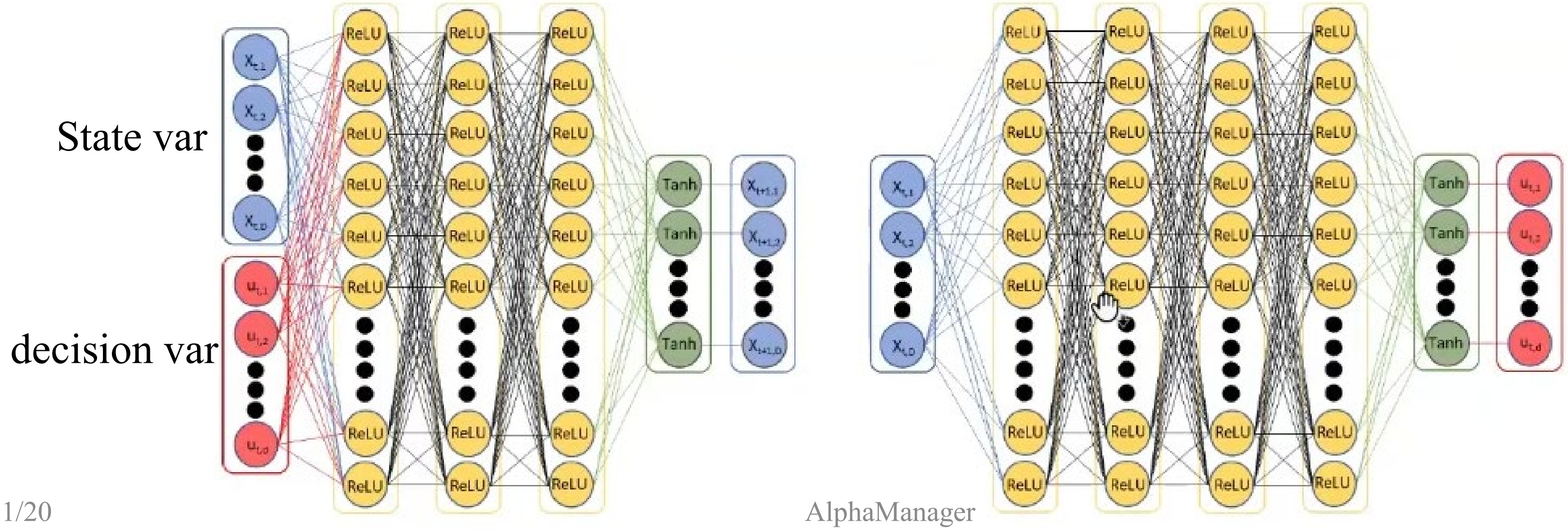
AlphaManager architecture

Predictive Environment module (PEM, 10 auxiliaries, 300+64+5):

- Warm-up 1976Q1 - 1991Q4; quarterly rolling training & test.

AlphaManager Policy Module: RL 256:64+5:

$$\max_{\{u_{t_0},\ldots,u_{t_0+T}\}} \mathbb{E}_{t_0} \sum_{t=t_0}^{t_0+T} r(X_t, u_t) \qquad s.t.\ \Delta X_{t+1} = f(X_t, u_t) + \varepsilon_{t+1}$$



Panel A: Predictive Environment Module

Panel B: Policy Module

Predictive Environment Module

$$\Delta X_{t+1} = f(X_t, u_t) + \varepsilon_{t+1},\ \varepsilon_{t+1} \sim N(0, \Sigma)$$

- X_t state variable: internal & external
- u_t managerial actions
- ΔX_{t+1} corresponds to current state and managerial actions
- Out-of-sample prediction
  - Overfitting: add penalty
  - Data shift: model ambiguity

Reinforcement Learning Module

$$\max_{\{u_{t_0},\ldots,u_{t_0+T}\}} \mathbb{E}_{t_0} \sum_{t=t_0}^{t_0+T} r(X_t, u_t) \qquad s.t.\ \Delta X_{t+1} = f(X_t, u_t) + \varepsilon_{t+1},\ \varepsilon_{t+1} \sim N(0, \Sigma),$$

- $r(X_t, u_t)$ reward function
- Assume decisions are contingent on the state variable

$$u_t = g(X_t),$$

- $\rightarrow$

$$\begin{aligned} \max_{g(\cdot)} \quad & \mathbb{E}_{t_0} \sum_{t=t_0}^{t_0+T} r(X_t, u_t) \\ s.t. \quad & \Delta X_{t+1} = f(X_t, u_t) + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim N(0, \Sigma) \\ & u_t = g(X_t). \end{aligned}$$

Robust Control and Ambiguity

- Ambiguity: dispersion in the predictions across models
- A bag of PEMs, indexed by i= 1,2, ... I, and ambiguity aversion
  - Mapping (X_t, u_t) to $\hat{X}_{t+1}^{i}$
- Boosting error: the greatest dispersion among model predictions

$$\text{BoostingError}(X_t, u_t) = \frac{1}{D}\sum_{d=1}^{D}\left(\max_{i=1,2,\ldots,I}\hat{X}_{t+1,d}^{i} - \min_{i=1,2,\ldots,I}\hat{X}_{t+1,d}^{i}\right)^2,$$

- In Policy Module: objective function:

$$J(X_{t_0}) := \sum_{t=t_0}^{t_0+T}\min_{i=1,2,\ldots,I} r^i\left(X_t, g(X_t)\right) - \delta\cdot\max\left\{0, \max_{t=t_0,\ldots,t_0+T}\text{BoostingError}\left(X_t, g(X_t)\right) - \theta(X_{t_0})\right\}$$

Data

- Data: Compustat (firm fundamentals), CRSP (market return andvolatility), and Chicago Fed (macro state variables)
- From 1976 to 2023, quarterly
  - 20,485 different firms ranging from1976:Q1 to 2023:Q4, with 784,460 firm-quarter observations
- State variables
  - 12 fundamental + 2 market + 4 macro (+lag + growth)
- Decision variables
  - 9 dimensions of actions (+growth)
- **PEM: input 268 → output 14 (12 fundamental change + ret/vol change)**
- **RLM: input 250 (268 - 18)→ output 9 (change)**
- Baseline objective: mktcap/enterprise value growth in short/long horizon

Empirical results

- PEM's Predictions of Firm Outcomes

	Ignoring Control		With Control	
State Variable	Training R^2	Test R^2	Training R^2	Test R^2
Book Asset Growth	-9.39%	-12.39%	61.28%	62.65%
Current Asset Growth	-2.16%	-5.94%	53.34%	57.01%
Gross Revenue Growth	36.43%	32.35%	38.76%	35.10%
Accounts Payable Growth	26.84%	26.17%	30.50%	30.02%
COGS Growth	34.95%	30.92%	36.45%	32.78%
Net Interest Paid Growth	81.36%	82.77%	81.58%	82.97%
Inventory Growth	16.48%	13.78%	21.08%	18.17%
Current Liability Growth	10.59%	6.93%	28.30%	25.61%
Receivables Growth	24.95%	22.86%	30.30%	28.33%
Net Income Growth	36.89%	32.88%	39.19%	35.58%
Trading Volume Growth	18.92%	15.90%	21.99%	19.92%
Log Gross Return Growth	49.86%	42.73%	52.34%	45.99%
Market Cap Growth	3.26%	-9.62%	12.98%	2.42%
Enterprise Value Growth	-2.29%	-12.83%	19.40%	10.57%

- Controls more important for some state evolution.

Empirical results

- Heterogeneous PEM performance

variable		full sample		pre-dotcom		dotcom-GFC		post-GFC	
		mean	std	mean	std	mean	std	mean	std
log_book_asset	low	2.95%	9.34%	3.33%	9.89%	2.96%	9.07%	2.66%	9.04%
	high	1.60%	6.28%	1.82%	6.58%	1.79%	6.70%	1.32%	5.75%
log_cogs	low	4.12%	12.40%	4.18%	12.10%	4.01%	12.25%	4.14%	12.70%
	high	2.75%	10.55%	3.09%	11.64%	2.66%	10.25%	2.53%	9.82%
log_current_liabilities	low	6.55%	15.11%	6.52%	14.43%	6.40%	14.87%	6.66%	15.74%
	high	5.23%	13.24%	5.65%	13.76%	5.46%	13.75%	4.78%	12.50%
log_market_cap	low	12.53%	21.42%	12.99%	21.63%	13.62%	22.77%	11.55%	20.38%
	high	7.68%	15.70%	9.92%	18.70%	8.30%	16.43%	5.55%	11.96%
log_enterprise_val	low	10.34%	18.96%	11.53%	20.69%	11.55%	20.10%	8.73%	16.65%
	high	6.06%	13.67%	8.59%	17.98%	6.14%	12.80%	4.05%	9.24%
macro1 (risk)	low	4.91%	5.78%	4.62%	5.33%	4.63%	5.37%	4.97%	6.11%
	high	6.00%	6.71%	6.42%	6.80%	5.99%	6.54%	4.84%	5.88%

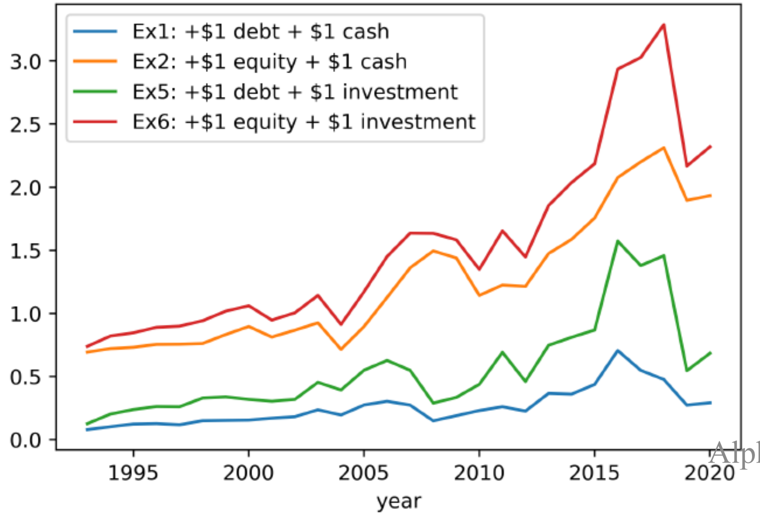
- Subsample episodes: pre-dot com, dot com to GFC, post-GFC
- Book asset: small firms has higher prediction error and std, pre-dotcom has the highest mean and std
- COGS: both higher and lower halves have declining average MSE
- Market cap, enterprise value, and macr1 (risk): highest prediction error during dotcom to GFC

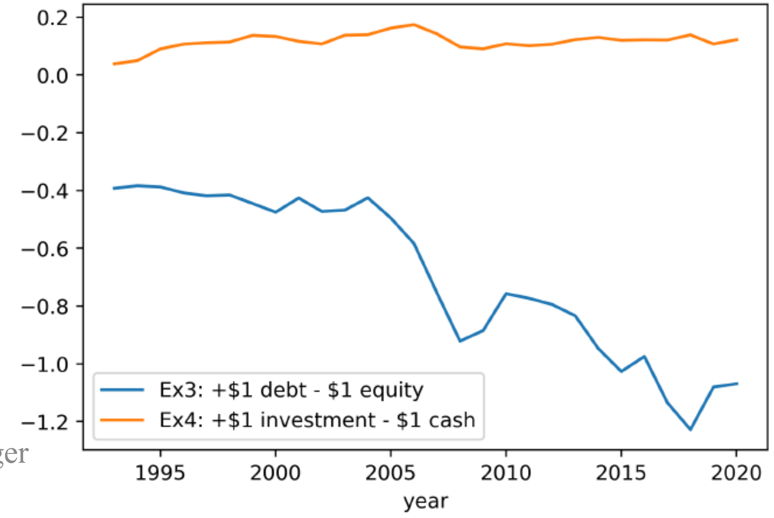
Empirical results

- PEM Application: Recapitalization Analysis

How does enterprise value change if a firm:

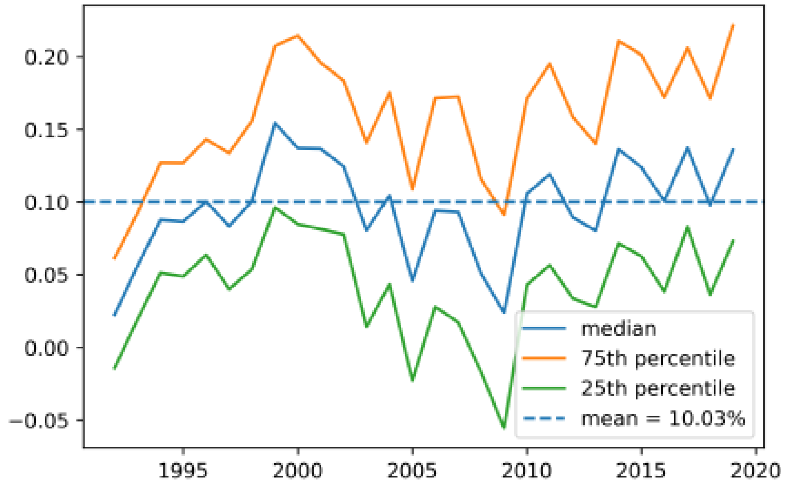
1. raises $1 more debt and put that $1 into its cash savings
2. raises $1 more equity and put that $1 into its cash savings
3. raises $1 more debt and $1 less equity
4. puts $1 cash into investment
5. raises $1 more debt and put that $1 into investment
6. raises $1 more equity and put that $1 into investment

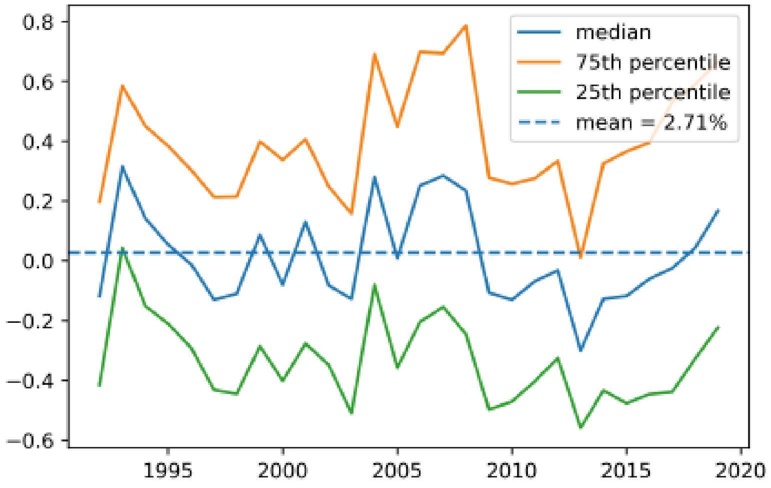




Optimal policy

Figure 6: Out-performance of AlphaManager with Short-term Objectives

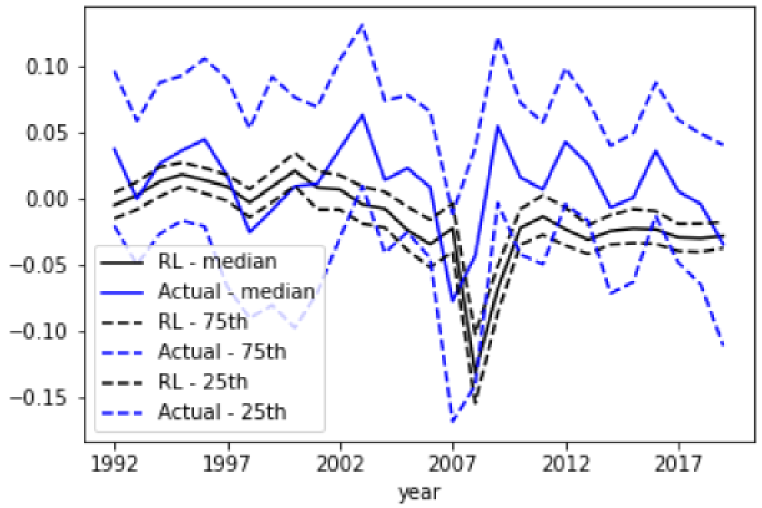


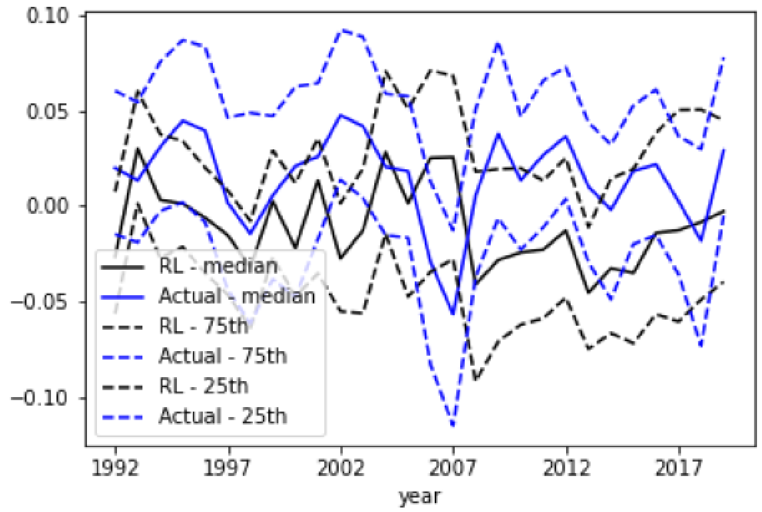


- Mktcap: 10.03%; enterprise value:2.71%
- Great dispersion

Optimal policy

- Long-term performance under short-termist





  - historical actions may generate better long-term performance

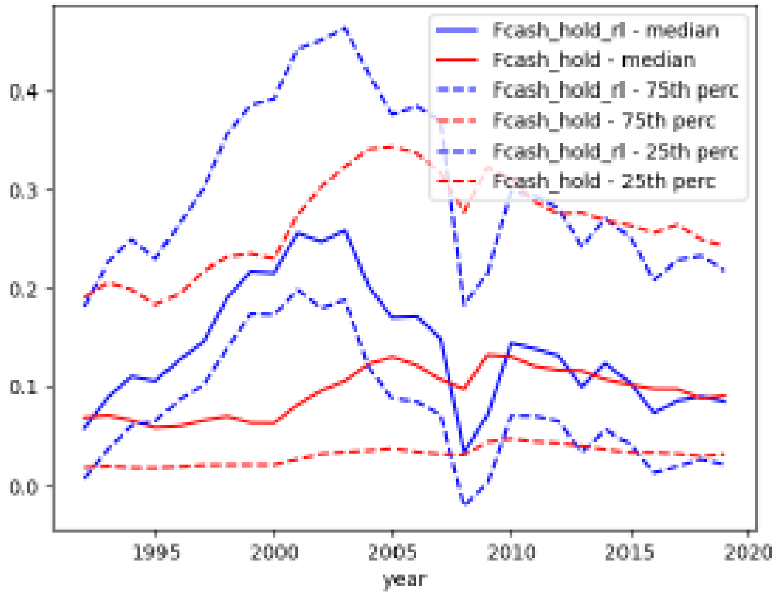
Optimal policy

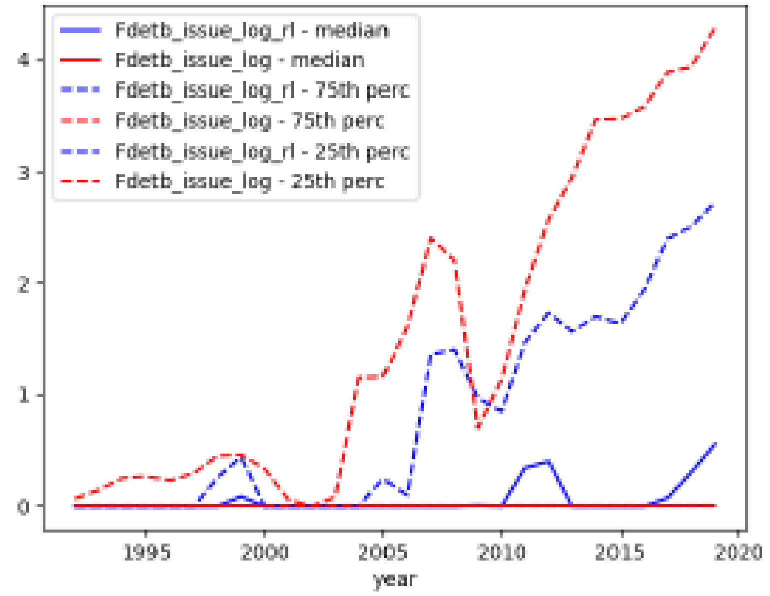
- Long-term oriented
  - Mktcap: 8.73% (3.09% in the 1st quarter)
  - Enterprise value: 4.43% (1.28% in the 1st quarter)

Table 9: Heterogeneous Performance of AlphaManager Across Firms

variable		full sample		pre-dotcom		dotcom-GFC		post-GFC	
		mean	std	mean	std	mean	std	mean	std
log_at	low	8.11%	3.73%	8.13%	4.00%	6.48%	4.74%	6.67%	4.92%
	high	1.99%	4.62%	4.54%	4.02%	0.74%	5.07%	0.17%	3.50%
log_act	low	10.13%	4.73%	6.75%	4.35%	5.14%	5.63%	4.96%	6.41%
	high	3.16%	4.68%	5.30%	4.29%	2.18%	5.38%	2.04%	3.74%
log_sale	low	8.80%	4.24%	7.98%	3.87%	6.18%	4.88%	6.08%	5.48%
	high	2.25%	4.68%	4.37%	4.11%	1.15%	5.31%	1.00%	3.93%
log_ap	low	8.40%	3.96%	7.66%	4.30%	5.69%	5.29%	5.97%	5.45%
	high	2.52%	4.69%	4.83%	4.05%	1.43%	5.27%	0.82%	3.81%
log_cogs	low	8.64%	4.06%	7.81%	3.89%	6.02%	4.90%	5.82%	5.56%

Optimal vs. empirical





- Similar trend, but different

Conclusion

- They build a predictive environment module through supervised neural networks, and then add a policy module through deep reinforcement learning.
- This framework not only better explains and predicts corporate outcomes in- and out-of-sample, but also identifies important managerial decisions while offering effective dynamic policies

Extension

- Manager behavior characteristics
- Combined objective function
- New hypothesis generation
- Market risk premium prediction with ambiguity control