University of Luxembourg


A Sustainable and Trustworthy AI Recommitment System (STAIRS)

As part of the STAREBEI project: “Toward A.I. Recommitment Strategies for ESG integration in Private Equity”
Content

Context & motivation

Evolutionary Learning of Private Equity Recommitment Strategies

Proximal Policy Optimisation for a Private Equity Recommitment System

Conclusion & Perspectives
Private Equity (PE)

- Alternative investment class
- Has gained a great amount of influence in today’s financial marketplace
- Included in the portfolio of sovereign wealth funds, pension funds …

Source: Federal Reserve Bank of Dallas, IFSL, EVCA/Thomson Venture Economics/PricewaterhouseCoopers
Private Equity Funds

Private Equity Fund – Cash Flow Model

**Investment Stage**
Year 1 through year 4-5, typically
- Capital is committed and drawn down
- Investments are made in portfolio companies

**Development stage**
Year 3 to year 8, typically
- Initial investment starts to mature
- Mature investments are exited
- Cash distributions are paid to investors
- Follow-on investments are made

**Maturity/Liquidation stage**
Year 8+, typically
- Most investments have been exited
- Several investments are left to “wind down”

Investing directly to companies requires high level of expertise, experience and staff incentives.

Institutional investors prefers to invest as Limited Partners.

LPs commit capital to the fund. General Partner (GP) calls the committed capital.
PE characteristics

- Stakes in PE are illiquid due to restriction on sales
- Exposure to PE by investing in new funds in which they commit
- Capital is drawn down gradually over several years
- Very often Capital is not entirely called
- Payouts (distributions) occurrence vary between funds
- Most of these distributions cannot be reinvested immediately and are recommitted to new PE funds

Consequently:
- Cash inflows and outflows are uncertain
- Investor have no control
- Can lead to PE misallocation
How to maintain high PE allocation

- **Underinvestment** because of undrawn Capital may lead to a drop of portfolio performance

- **Overinvestment** due to too large commitments may result in a liquidity shortfall

- Find a trade-off by keeping investment degree close to 1:
  - \[ ID_t = \frac{NAV_t}{NAV_t + Cash_t} \approx 1 \text{ for all period } t \]

- A multi-period portfolio optimization

- Dynamic evolution of PE portfolio

- Need a strategy to be applied at each period t
Solve multiple single-period portfolios

- Based on single-period portfolio optimization problem for each period $t$

- $\min_{C_t} E_t (1 - ID_{t+1})^2$ with $E_t$ the conditional expectation at end of period $t$

- Analytical solution found at $C_t = E_t \left( \frac{Cash_t + D_{t+1} - \sum \gamma_{t+1,i+1} C_{t-1}}{\gamma_{t+1,1}} \right)$

- Involve data from period $t+1$

  with $\gamma_{t+1,t-i}$, the fraction of capital committed $i$ periods ago and called at $t + 1$
Recommitment rules (deZwart 2012)

- “Private Equity Recommitment Strategies for Institutional Investors”
- Propose for the Dutch Pension Fund (APG)
- No cashflow forecasting
- Manually designed rules of thumb as strategies:
  
  - $DZ^1$: $C_t = D_t$
  
  - $DZ^2$: $C_t = D_t + UC_{t-p}$
  
  - $DZ^3$: $C_t = \frac{1}{ID_t} (D_t + UC_{t-p})$

- Can we find better ones automatically? Can we learn to optimize strategies?
- With additional constraints, manually designed rules become unsuitable
Evolutionary Learning of Private Equity Recommitment Strategies
Genetic Programming

- Search heuristic that is inspired by Charles Darwin's theory of natural evolution

- “Individuals with traits that enable them to adapt to their environments will help them survive and have more offspring, which will inherit those traits.”

- Technique of evolving programs

- Global Optimization approach:
  - Derivative-free
  - No assumption

- Evolving programs already mentioned by A. Turing (1950’s)
Recommitment strategies are programs

- Programs have traits that can be evolved

- Why Evolutionary learning?
  - Learning is an optimization problem
  - Learning $\Leftrightarrow$ Recognizing

- “Strategies with traits that enable them to improve the Investment Degree will help them survive and have more offspring, which will inherit those traits.”
How do we evolve program?

- Program ↔ Hierarchical Data structure
- Abstract Syntax Tree (AST)
- Two main operators:
  - Crossover -- exploitation
  - Mutation – exploration
- The best individual will survive?
- How do you measure it?
How to measure the fitness?

\[ C_t = \frac{1}{ID_t} (D_t + UC_{t−p}) \]

**Program**

**Function**

**Recommitment Strategy**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>Add two inputs</td>
</tr>
<tr>
<td>-</td>
<td>Subtract two inputs</td>
</tr>
<tr>
<td>*</td>
<td>Multiply two inputs</td>
</tr>
<tr>
<td>%</td>
<td>Divide two inputs with protection</td>
</tr>
<tr>
<td>min</td>
<td>Minimum b.t.w. two inputs</td>
</tr>
<tr>
<td>max</td>
<td>Maximum b.t.w. two inputs</td>
</tr>
<tr>
<td>( C_t )</td>
<td>Contributions at ( t )</td>
</tr>
<tr>
<td>( D_t )</td>
<td>Distributions at ( t )</td>
</tr>
<tr>
<td>( ID_t )</td>
<td>Investment degree at ( t )</td>
</tr>
<tr>
<td>( NAV_t )</td>
<td>Net Asset Value at ( t )</td>
</tr>
<tr>
<td>( error_t )</td>
<td>Deviation to ideal ID at ( t )</td>
</tr>
<tr>
<td>( DZ^3(t) )</td>
<td>deZwart’s strategy n°3 [8] at ( t )</td>
</tr>
<tr>
<td>( UC_{t−24} )</td>
<td>Uncalled capital for commitments made 24 quarters ago</td>
</tr>
<tr>
<td>( Ccommit_{t−24} )</td>
<td>Capital committed for 24 quarters</td>
</tr>
</tbody>
</table>
How to measure the fitness?

\[ \text{obj} = \int_{t_1}^{t} |1.0 - UCB(t)| \cdot dt + K \cdot (t_2 - t) \]

with \( UCB(t) = E_p(ID_t) + 2\sigma_p(ID_t) \)
Experimental setup

- Artificial cashflows:
  - PE players protect their rich data histories
- Private market data providers generally sell data
- Cover very specific periods and incomplete
- Synthetic cashflows generated from a stochastic version of the Yale Model

Genetic Programming parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Runs</td>
<td>30</td>
</tr>
<tr>
<td>Generations</td>
<td>50</td>
</tr>
<tr>
<td>Population size</td>
<td>500</td>
</tr>
<tr>
<td>Crossover Probability (CXPB)</td>
<td>0.85</td>
</tr>
<tr>
<td>Mutation Probability (MUTPB)</td>
<td>0.1</td>
</tr>
<tr>
<td>Reproduction Probability</td>
<td>0.05</td>
</tr>
<tr>
<td>Tree initialization method</td>
<td>Ramped half-and-half</td>
</tr>
<tr>
<td>Selection Method</td>
<td>Tournament selection with size=7</td>
</tr>
<tr>
<td>Depth limitation</td>
<td>17</td>
</tr>
<tr>
<td>Crossover Operator (CX)</td>
<td>One crossover point</td>
</tr>
<tr>
<td>Mutation Operator (MUT)</td>
<td>Grow</td>
</tr>
</tbody>
</table>

Simulation parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cashflows frequency</td>
<td>quarterly</td>
<td>quarterly</td>
</tr>
<tr>
<td>Investment period</td>
<td>26 years</td>
<td>26 years</td>
</tr>
<tr>
<td>Funds per recommitment</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Fund selection</td>
<td>ESG score</td>
<td>ESG score</td>
</tr>
<tr>
<td>Number of simulated portfolios (per evaluation)</td>
<td>250</td>
<td>1000</td>
</tr>
<tr>
<td>Distributed simulation</td>
<td>True</td>
<td>False</td>
</tr>
</tbody>
</table>
Experimental results
Experimental results

![Graph showing experimental results over years]
Experimental results

Best strategy obtained from the 30 runs, i.e., $S^{best}(t) =$
\[
\max (-\text{Cash}_t \times D_t + D Z^3(t), \min (\text{Cash}_t, D_t + 2U C_{t-24})) + \\
\min (\text{Cash}_t, \max (D^2_t, D_t + 2U C_{t24}))
\]
Proximal Policy Optimisation for a Private Equity Recommitment System
Learning Recommitment policies

- Using a policy-based algorithm ~> Proximal Policy Optimization (PPO)

- Target recommitment policies maintaining an Investment Degree close to 1

- Policy-based VS Value-based:
  - Avoid computational burden to compute all state-values
  - Action space is continuous

- **Drawbacks:**
  - On-policy approaches
  - Large number of simulations
RL model of the PE recommitment problem

Agent

Reward $r_t$

State $s_t$

Portfolio

Distributions

Capital calls

1.0

Recommitted capital

Action $a_t$
RL model of the PE recommitment problem

- State $s_t = < ID_t, D_t, CC_t, UC_{t-24}, Cash_t, NAV_t >$
  - Portfolio state
  - Important features to recommit

- Continuous action $a_t \Rightarrow$ capital recommitted into new PE funds

- Final reward $\sum_{t=1}^{T} ID_t \times 10^{\text{digits}(T) + 1} + \sum_{t=1}^{T} r_t^{\text{valid}}$
  - Global reward:
    - Based on ID
    - Only if no cash shortage
  - Local reward:
    - $r_t^{\text{valid}} = \begin{cases} 0 & \text{if } ID_t > 1 \\ 1 & \text{else} \end{cases}$

- Create a different order of magnitude between valid portfolios and invalid ones (constraint handling)

- Accumulated local reward + shifted global reward
Algorithm 1 PPO-clip version

1: Initialize policy parameters $\theta_1$ and value function parameters $\phi_1$
2: for $k \in \{1, ..., M\}$ do
3: Sample a set of trajectories $\{\tau_i\}_{i=1}^M$ using the policy $\pi_{\theta_k}$
4: Create a batch $\mathcal{B}$ of transitions $(s_t^i, a_t^i, r_t^i) \ \forall t \in \{1, ..., |\tau_i|\} \ \forall i \in \{1, ..., \ M\}$
5: Compute rewards-to-go $\hat{R}_t^i$, i.e. rewards from action $a_t^i$, $\forall t \in \{1, ..., |\tau_i|\} \ \forall i \in \{1, ..., \ M\}$
6: Estimate the advantages $A^{\pi_{\theta_k}}(s_t^i, a_t^i)$ using the value function $V_{\phi_k}$
7: Perform policy update:
   
   $\theta_{k+1} = \arg \max_{\theta} \frac{1}{M} \sum_{i=1}^M \frac{1}{|\tau_i|} \sum_{t=1}^{T_i} \left[ \min \left( A^{\pi_{\theta}}(s_t^i, a_t^i), \frac{\pi_{\theta}(a_t^i|s_t^i)}{\pi_{\theta_{old}}(a_t^i|s_t^i)}, g(\epsilon, A^{\pi_{\theta}}(s_t^i, a_t^i)) \right) \right]$
   
   with $g(\epsilon, A^{\pi_{\theta}}(s_t^i, a_t^i)) = \text{clip} \left( \frac{\pi_{\theta}(a_t^i|s_t^i)}{\pi_{\theta_{old}}(a_t^i|s_t^i)}, 1 - \epsilon, 1 + \epsilon \right)$
8: Perform value function update by minimizing mean-squared error:
   
   $\phi_{k+1} = \arg \min_{\phi} \frac{1}{M} \sum_{i=1}^M \frac{1}{|\tau_i|} \sum_{t=1}^{T_i} \left[ V_{\phi}(s_t^i) - \hat{R}_t^i \right]^2$
9: end for
Experimental setup

- **Artificial cashflows:**
  - PE players protect their rich data histories
  - Private market data providers generally sell data
  - Cover very specific periods and incomplete
  - Synthetic cashflows generated from a stochastic version of the Yale Model

- **Simulation parameters**

- **RL parameters**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>steps_per_epoch</td>
<td>26000</td>
</tr>
<tr>
<td>gamma</td>
<td>1</td>
</tr>
<tr>
<td>epochs</td>
<td>500</td>
</tr>
<tr>
<td># episodes</td>
<td>125000</td>
</tr>
<tr>
<td>clip_ratio $\epsilon$</td>
<td>0.2</td>
</tr>
<tr>
<td>$pi_{lr} / vf_{lr}$</td>
<td>$3 \times 10^{-4} / 1 \times 10^{-4}$</td>
</tr>
<tr>
<td>hidden layers</td>
<td>[64, 64]</td>
</tr>
</tbody>
</table>

- **Simulation parameters**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cashflows frequency</td>
<td>quarterly</td>
<td>quarterly</td>
</tr>
<tr>
<td>Investment period</td>
<td>26 years</td>
<td>26 years</td>
</tr>
<tr>
<td>Funds per recommitment</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Fund selection</td>
<td>ESG score</td>
<td>ESG score</td>
</tr>
<tr>
<td>Number of simulated portfolios</td>
<td>250</td>
<td>1000</td>
</tr>
<tr>
<td>(per evaluation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distributed simulation</td>
<td>True</td>
<td>False</td>
</tr>
</tbody>
</table>
Rewards evolution

![Graph showing the evolution of average rewards over epochs. The graph plots performance against epochs, with a peak around the middle and some fluctuations.](image-url)
Best policy obtained with PPO-clip
Comparison with existing strategies
Conclusion

- Strong influence in today’s financial marketplace

- PE challenges:
  - Stakes in PE are illiquid due to restriction on sales
  - Exposure to PE by investing in new funds in which they commit
  - Capital is drawn down gradually over several years
  - Very often Capital is not entirely called
  - Most of these distributions cannot be reinvested immediately and are recommitted to new PE funds

- Efficient recommitment policies/rules can be generated using Reinforcement Learning

- Next steps:
  - Multi-objective Reinforcement Learning
  - Multi-class assets portfolios
What’s next?
Providing a set of alternatives

- When liquidity is soft constraint
- Multi-class asset portfolios
Thank you for your attention