University of Luxembourg

E. Kieffer, F. Pinel, T. Meyer,
G. Gloukoviezoff, H. Lucius, P. Bouvry

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"

European Investment Bank

UNIVERSITÉ DU LUXEMBOURG

Content

Context & motivation

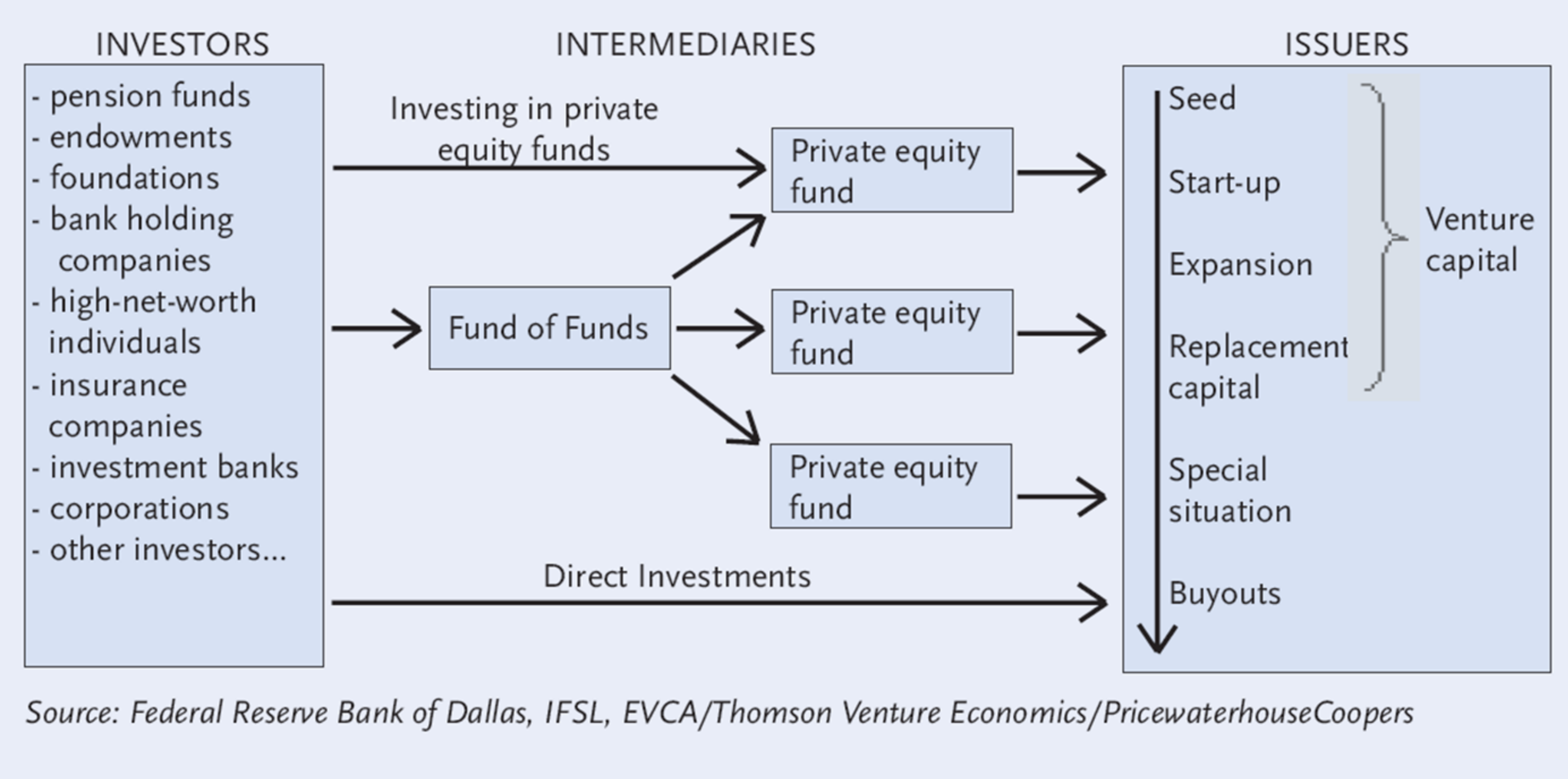
Evolutionary Learning of Private Equity Recommitment Strategies

Proximal Policy Optimisation for a Private Equity Recommitment System

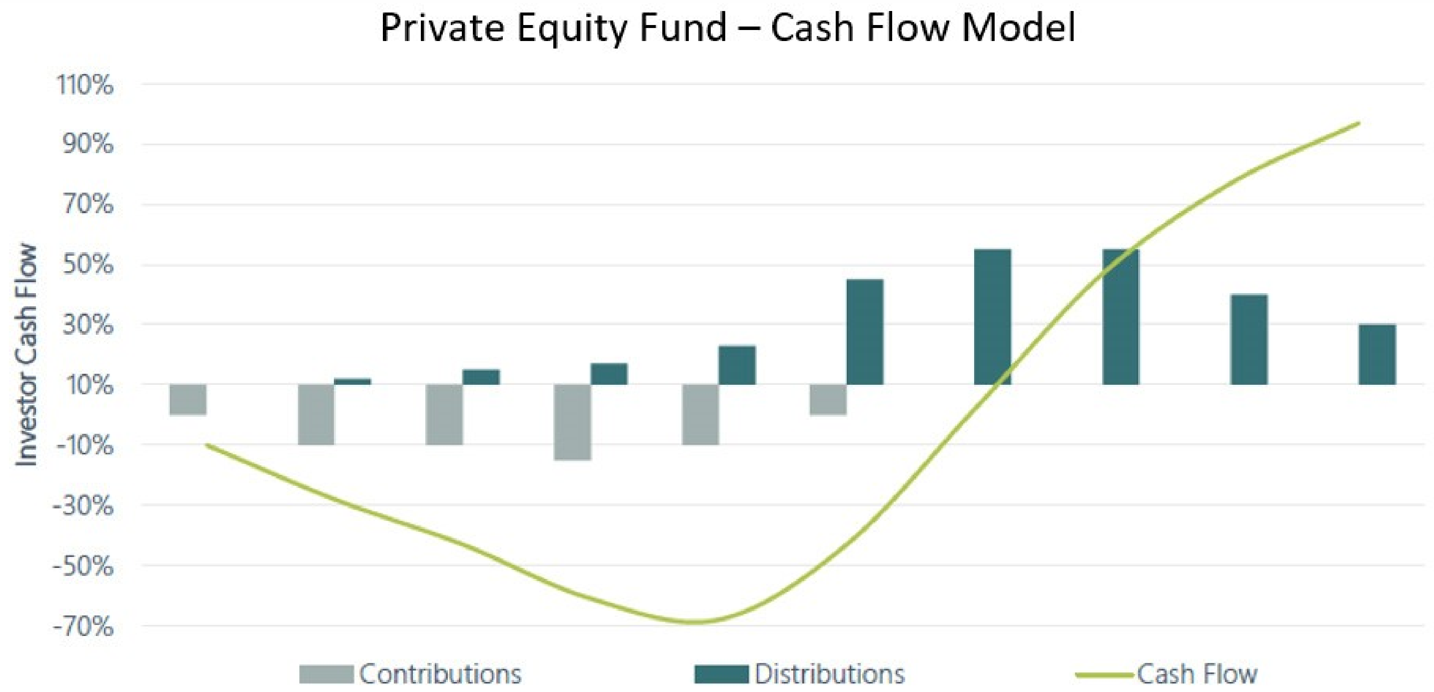
Conclusion & Perspectives

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 …



Private Equity Funds



Investment Stage	Development stage	Maturity/Liquidation stage
Year 1 through year 4-5, typically	Year 3 to year 8, typically	Year 8+, typically

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

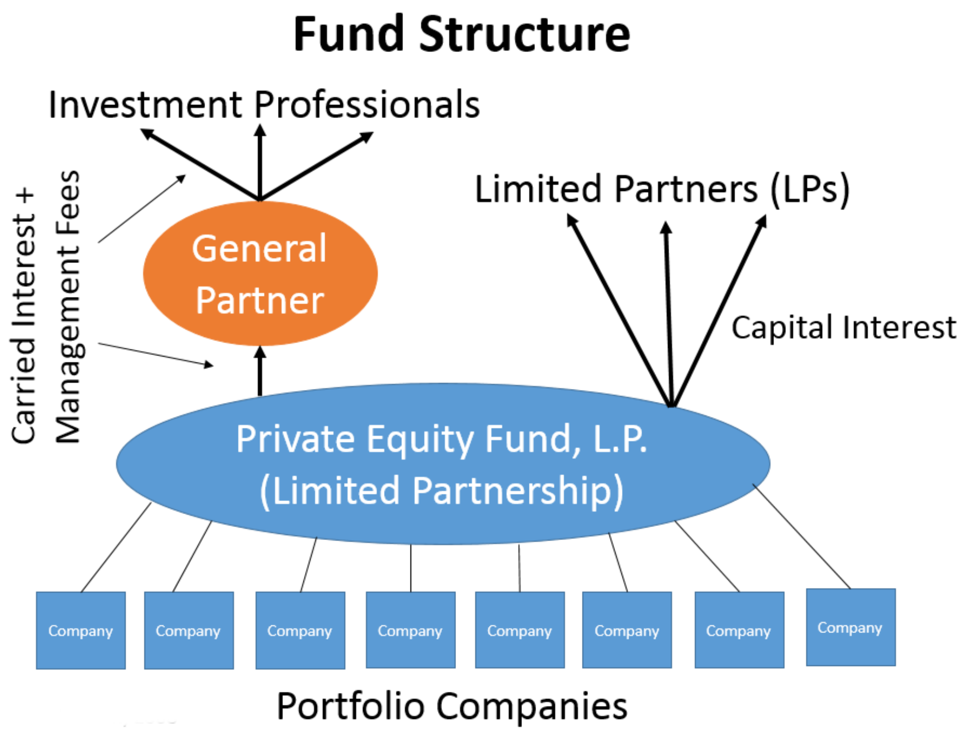
Maturity/Liquidation stage — Year 8+, typically

- Most investments have been exited
- Several investments are left to "wind down"

Schematic illustration. Source: RCP Advisors and Marquette Associates. As of 3/8/18.

Limited Partnership Funds

- 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 for all period t**

  - A multi-period portfolio optimization

- Dynamic evolution of PE portfolio

- Need a strategy to be applied at each period t

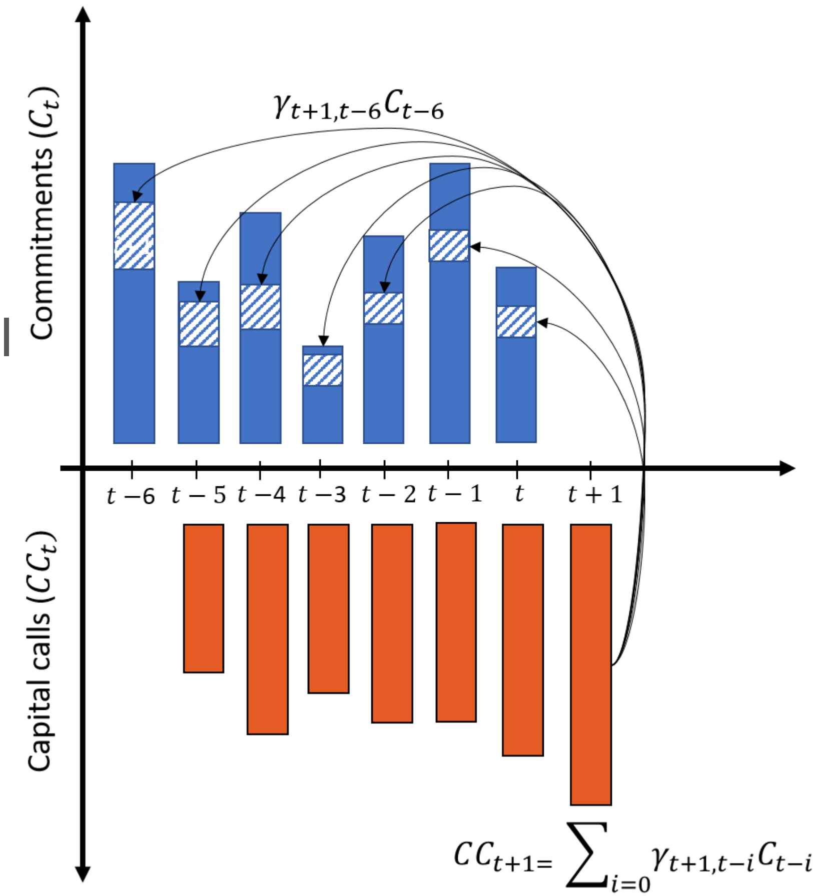
Solve multiple single-period portfolios

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

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

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

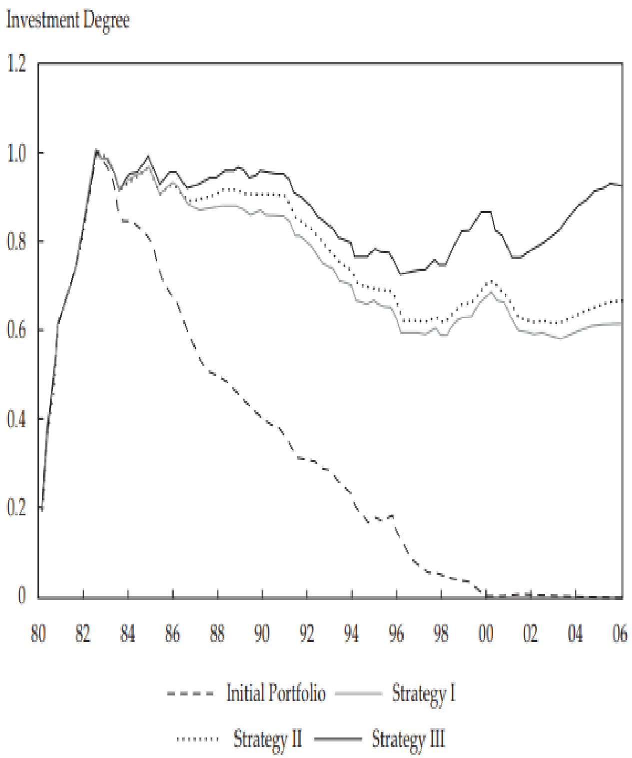
- Involve data from period t+1



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 : $C_t = \frac{1}{ID_t}(D_t + UC_{t-p})$



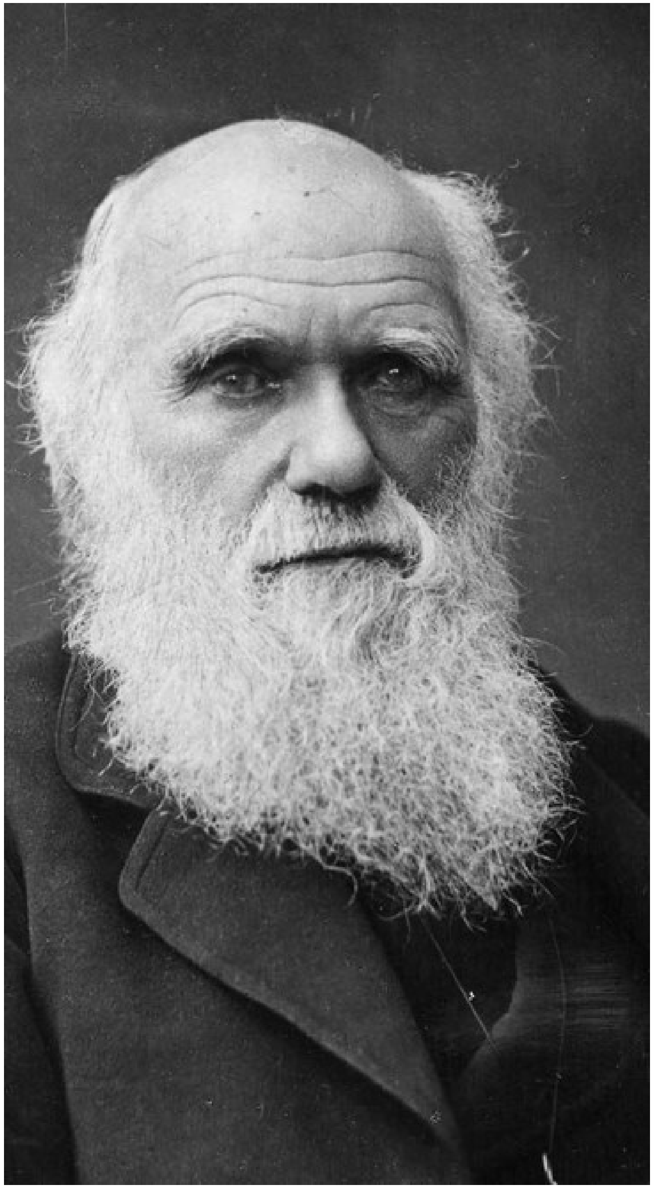
- Can we find better ones automatically ? Can we learn to optimize strategies ?
- With additional constraints, manually designed rules become unsuitable

Evolutionary Learning of Private Equity Recommitment Strategies

European Investment Bank

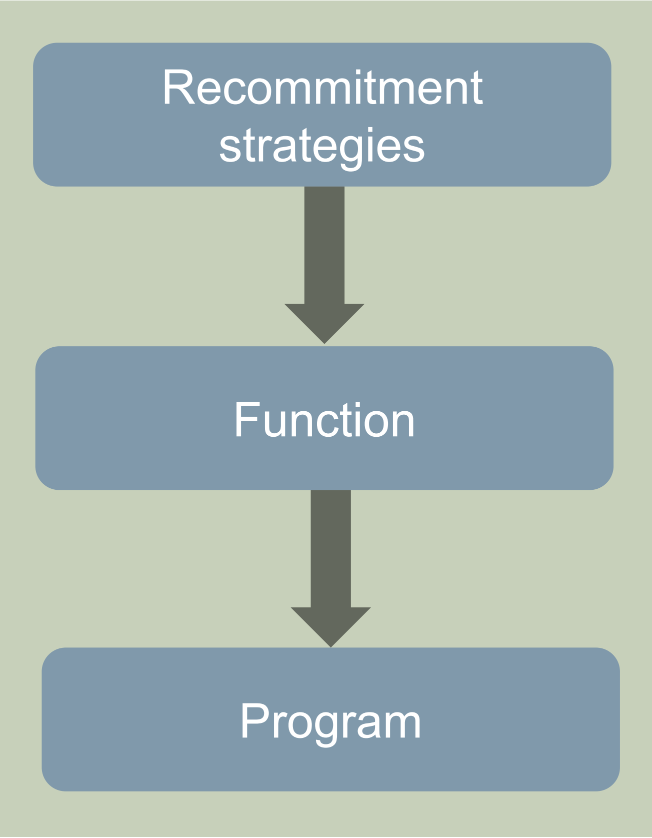
UNIVERSITÉ DU LUXEMBOURG

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)

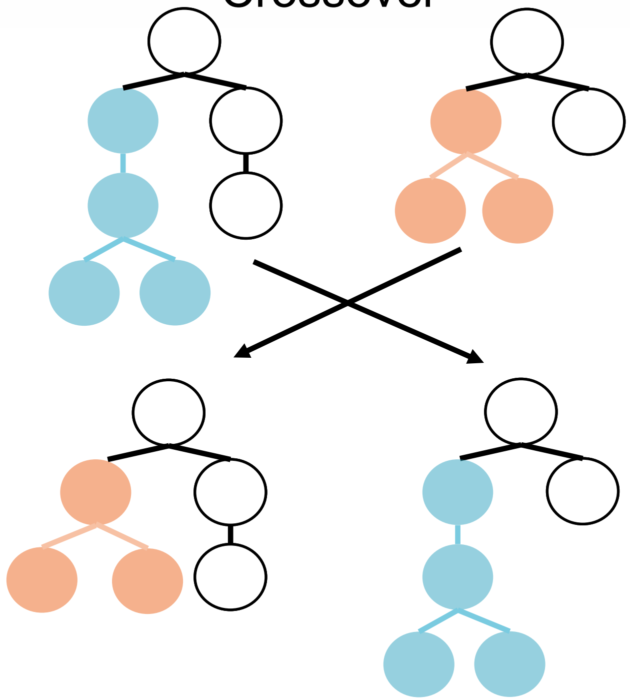
Evolving Strategies



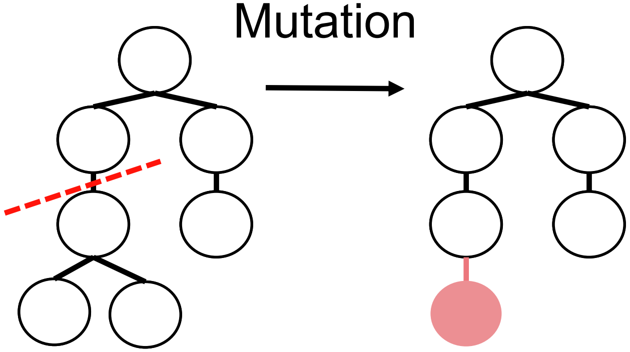
- Recommitment strategies are programs
- Programs have traits that can be evolved
- Why Evolutionary learning ?
  - Learning is an optimization problem
  - Learning ⇔ Recognizing
- “**Strategies** with **traits** that enable them to improve the Investment Degree will help them survive and **have more offspring**, which will inherit those **traits**.”

How do we evolve program ?

Crossover



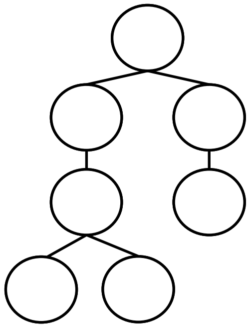
Mutation



- 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 ?

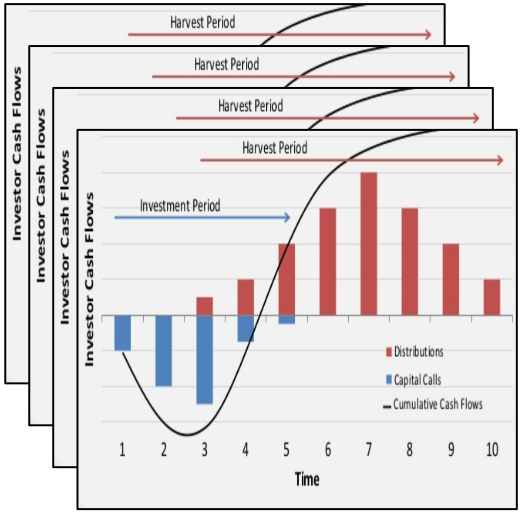
Program



Function

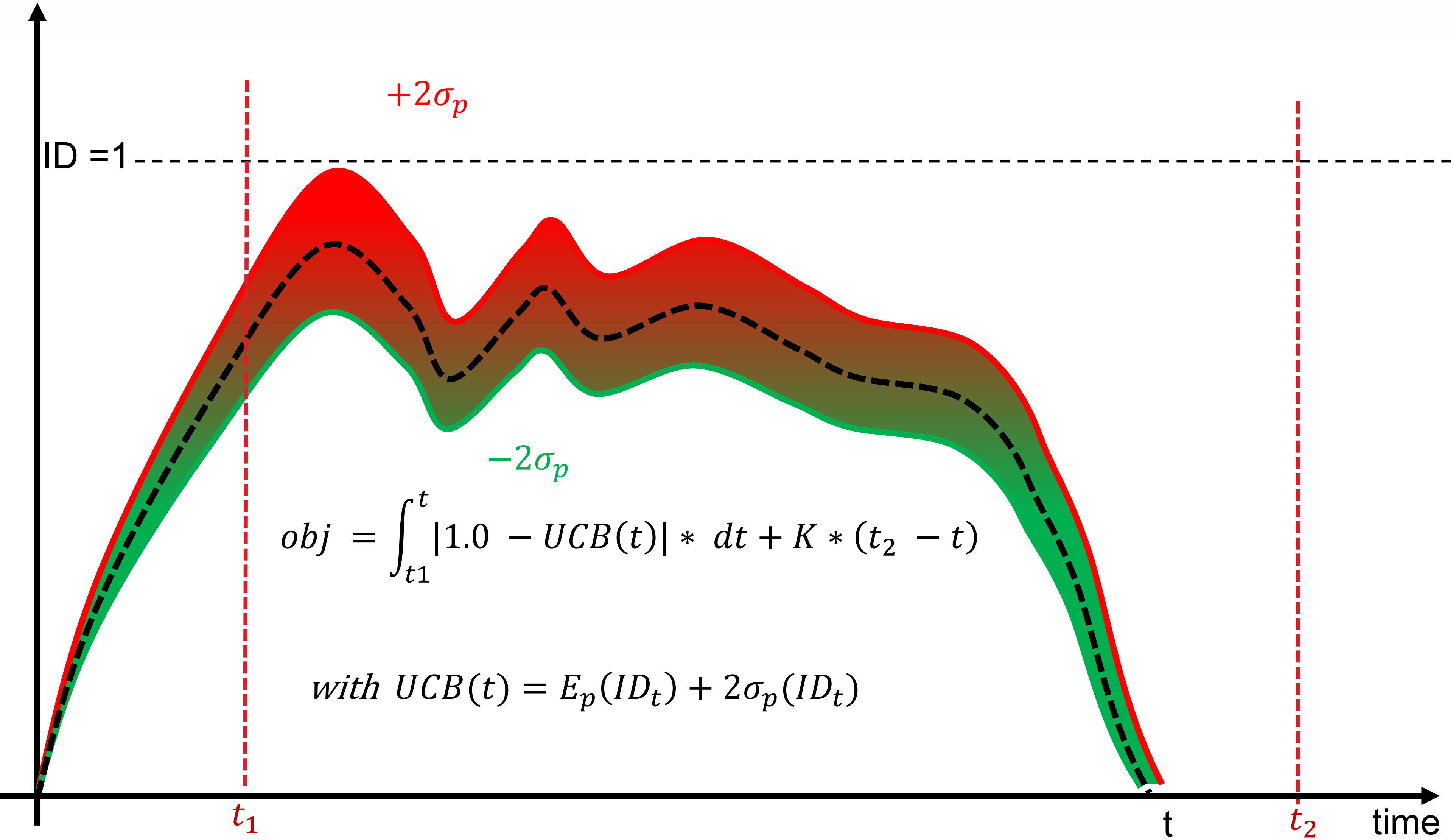
$$C_t = \frac{1}{ID_t}\left(D_t + UC_{t-p}\right)$$

Recommitment Strategy



Name	Description
	Operators
+	Add two inputs
-	Subtract two inputs
*****	Multiply two inputs
%	Divide two inputs with protection
min	Minimum b.t.w. two inputs
max	Maximum b.t.w. two inputs
	Terminal sets/ Arguments
C_t	Contributions at t
D_t	Distributions at t
ID_t	Investment degree at t
NAV_t	Net Asset Value at t
$error_t$	Deviation to ideal ID at t
$DZ^3(t)$	deZwart's strategy n°3 [8] at t
UC_{t-24}	Uncalled capital for commitments made 24 quarters ago
$CCommit_{t-24}$	Capital committed for 24 quarters

How to measure the fitness ?



$$obj = \int_{t1}^{t} |1.0 - UCB(t)| * dt + K * (t_2 - t)$$

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

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

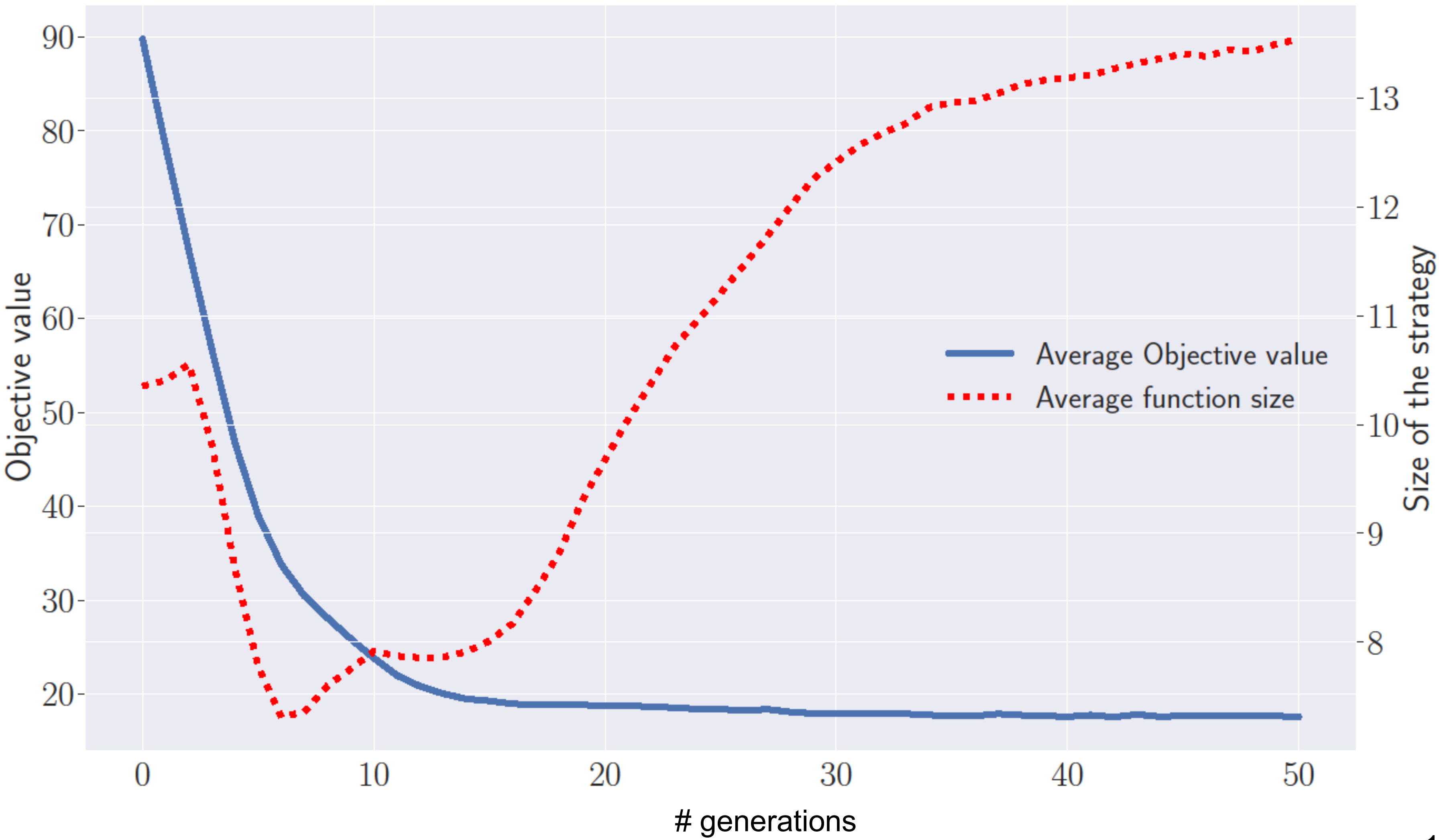
Genetic Programming parameters

Runs	30
Generations	50
Population size	500
Crossover Probability (CXPB)	0.85
Mutation Probability (MUTPB)	0.1
Reproduction Probability	0.05
Tree initialization method	Ramped half-and-half
Selection Method	Tournament selection with size=7
Depth limitation	17
Crossover Operator (CX)	One crossover point
Mutation Operator (MUT)	Grow

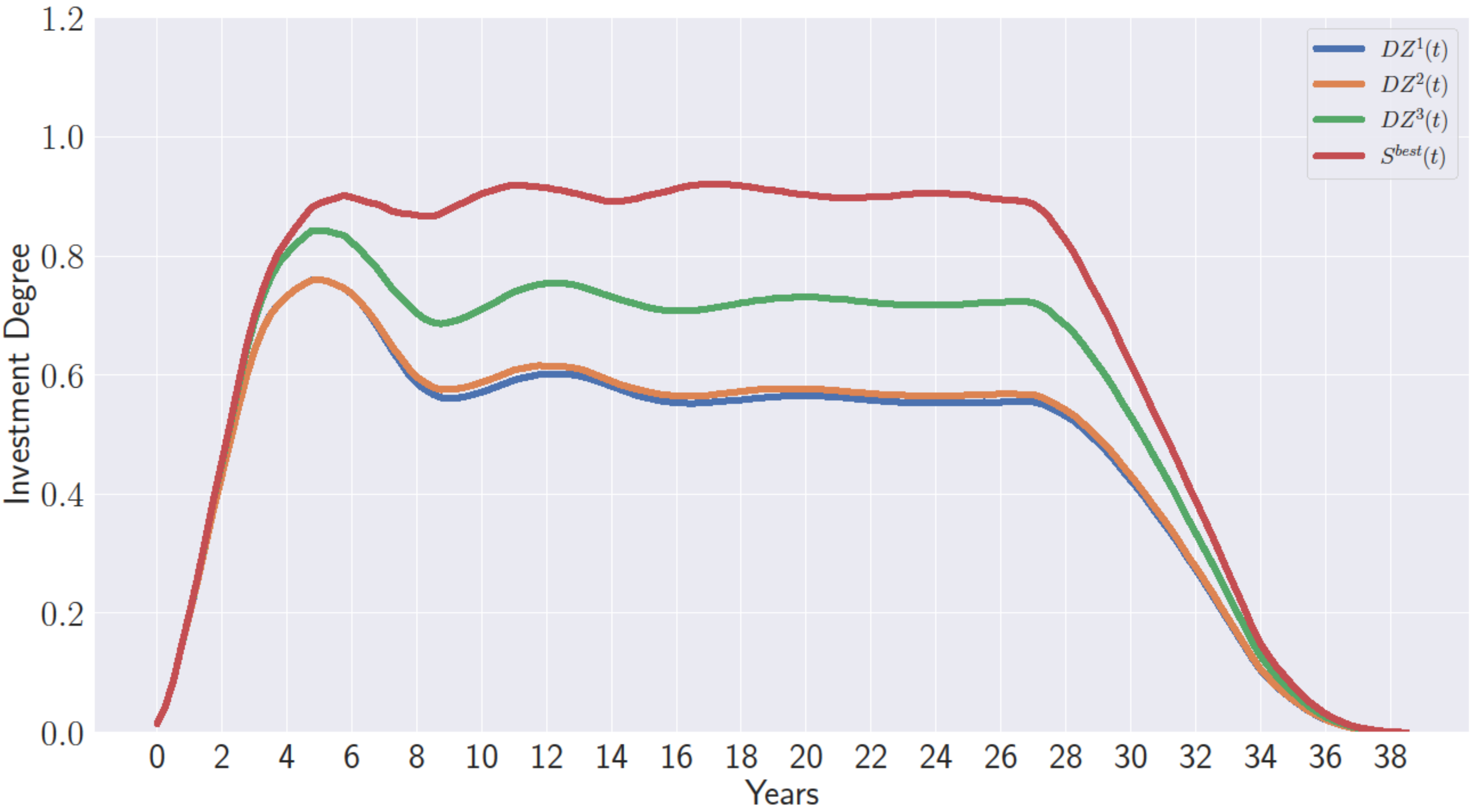
Simulation parameters

Parameters	**Training**	**Validation**
Cashflows frequency	quarterly	quarterly
Investment period	26 years	26 years
Funds per recommitment	4	4
Fund selection	ESG score	ESG score
Number of simulated portfolios (per evaluation)	250	1000
Distributed simulation	True	False

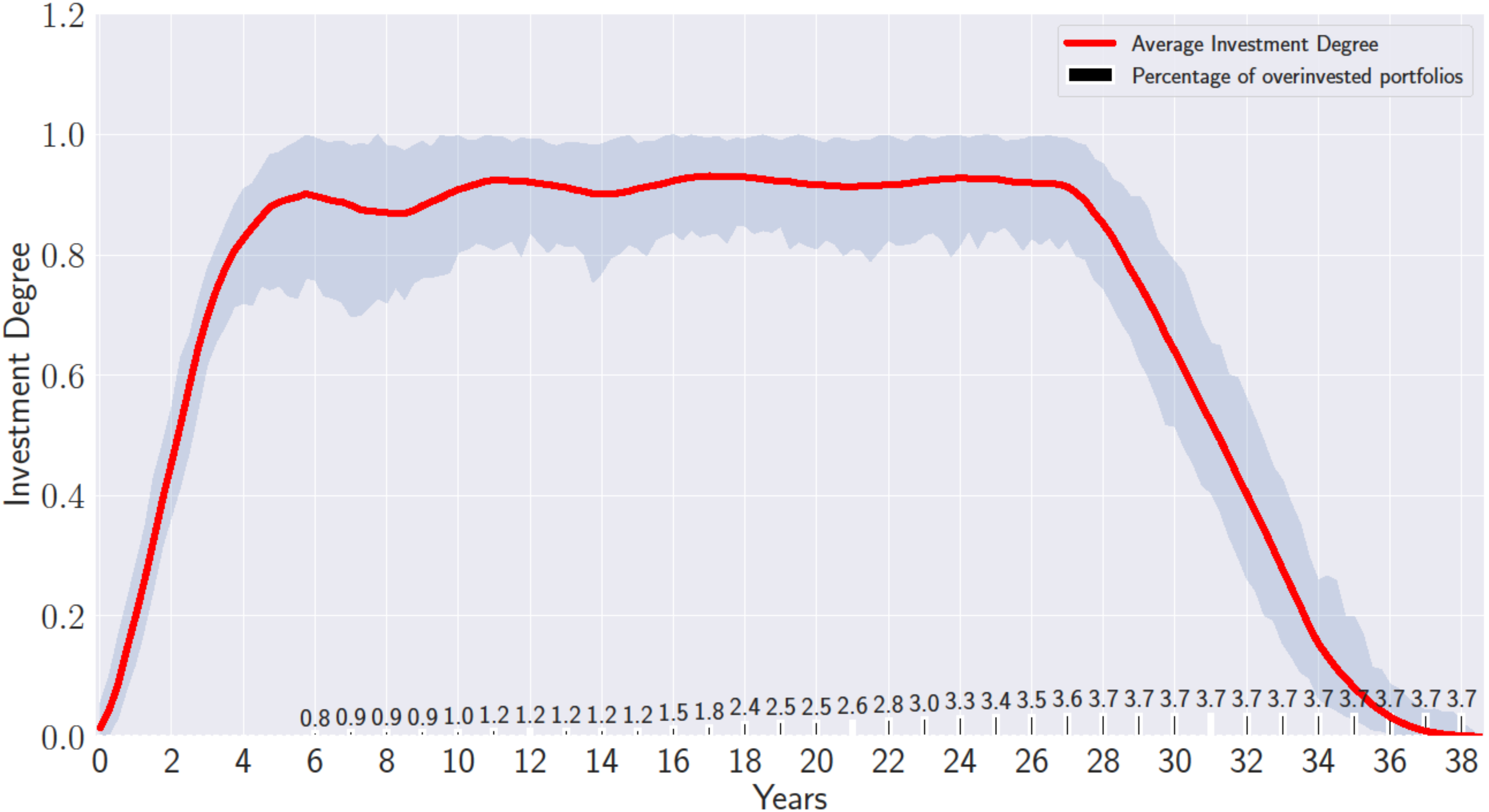
Experimental results



Experimental results



Experimental results



Best strategy obtained from the 30 runs, i.e., $S^{best}(t) = \max\left(-Cash_t \times D_t + DZ^3(t), \min\left(Cash_t, D_t + 2UC_{t-24}\right)\right) + \min\left(Cash_t, \max\left(D_t^2, D_t + 2UCt_{24}\right)\right)$

Proximal Policy Optimisation for a Private Equity Recommitment System

European Investment Bank

uni.lu UNIVERSITÉ DU LUXEMBOURG

Learning Recommitment policies

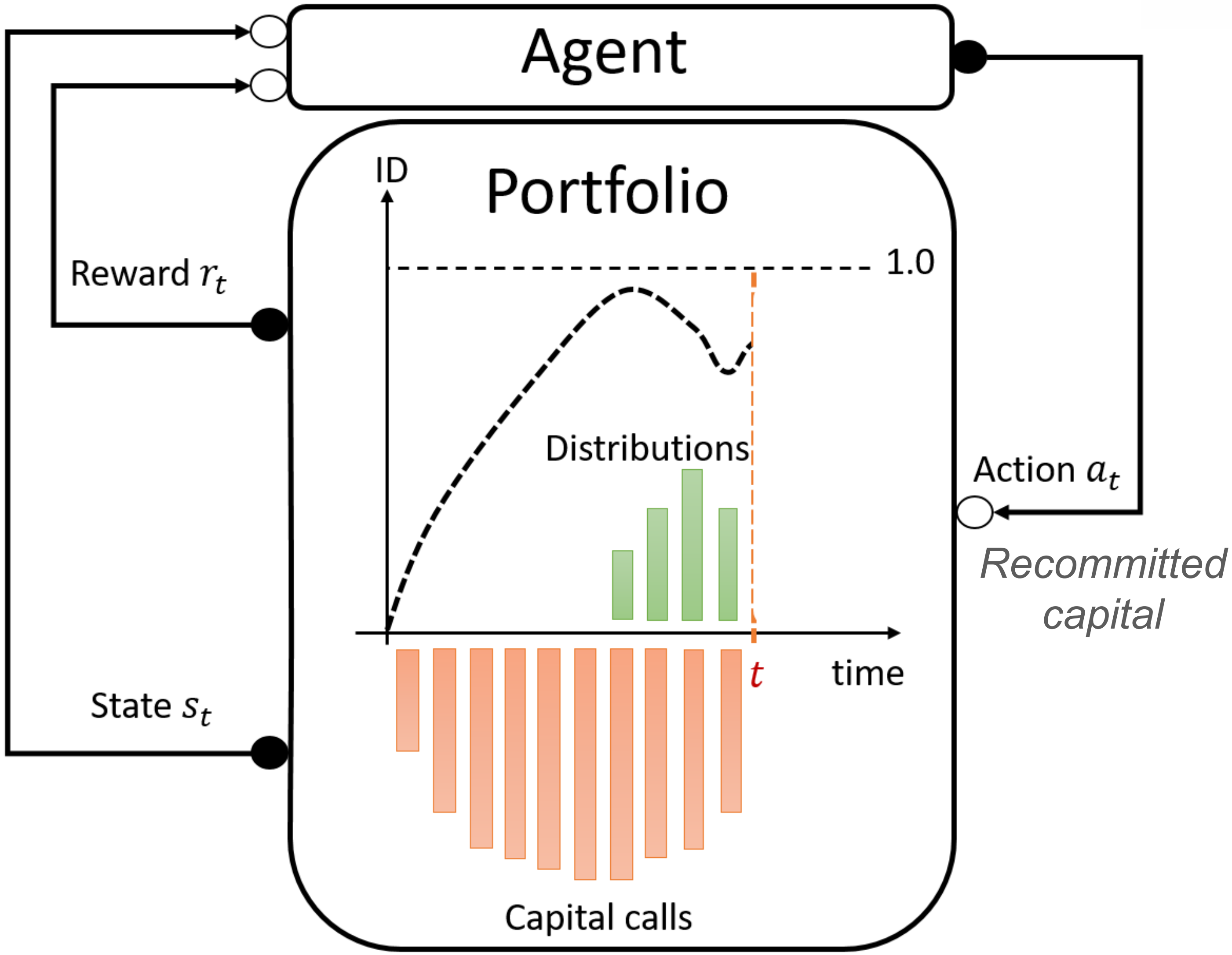
- Using a policy-based algorithm ~> Proximal Policy Optimization (PPO)

- 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



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 => capital recommitted into new PE funds

- Final reward $\underbrace{\sum_{t=1}^{T} ID_t \times 10^{(digits(T)+1)}}_{\text{Global reward}} + \underbrace{\sum_{t=1}^{T} r_t^{valid}}_{\text{Local reward}}$

Global reward:
- Based on ID
- Only if no cash shortage

Local reward:
- $r_t^{valid} = \begin{cases} 0 & if\ ID_t > 1 \\ 1 & else \end{cases}$

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

  - Accumulated local reward + shifted global reward

PPO-clip algorithm

Algorithm 1 PPO-clip version

Initialize policy parameters θ_1 and value function parameters ϕ_1
for $k \in \{1, \ldots, M\}$ **do**
    Sample a set of trajectories $\{\tau_i\}_{i=1}^{M}$ using the policy π_{θ_k}
    Create a batch $\mathcal{B}$ of transitions (s_t^i, a_t^i, r_t^i) $\forall t \in \{1, \ldots, |\tau_i|\}$ $\forall i \in \{1, \ldots, M\}$
    Compute rewards-to-go $\hat{\mathcal{R}}_t^i$, i.e. rewards from action a_t^i, $\forall t \in \{1, \ldots, |\tau_i|\}$ $\forall i \in \{1, \ldots, M\}$
    Estimate the advantages $A^{\pi_{\theta_k}}(s_t^i, a_t^i)$ using the value function V_{ϕ_k}
    Perform policy update:

$$\theta_{k+1} = \arg\max_{\theta} \frac{1}{M} \sum_{i=1}^{M} \frac{1}{|\tau_i|} \sum_{t=1}^{T_i} \left[\min \left(A^{\pi_\theta}(s_t^i, a_t^i) \frac{\pi_\theta(a_t^i | s_t^i)}{\pi_{\theta_{old}}(a_t^i | s_t^i)}, g\left(\epsilon, A^{\pi_\theta}(s_t^i, a_t^i)\right) \right) \right]$$

with $g\left(\epsilon, A^{\pi_\theta}(s_t^i, a_t^i)\right) = \text{clip}\left(\frac{\pi_\theta(a_t^i | s_t^i)}{\pi_{\theta_{old}}(a_t^i | s_t^i)}, 1 - \epsilon, 1 + \epsilon \right)$

    Perform value function update by minimizing mean-squared error:

$$\phi_{k+1} = \arg\min_{\phi} \frac{1}{M} \sum_{i=1}^{M} \frac{1}{|\tau_i|} \sum_{t=1}^{T_i} \left[V_\phi(s_t^i) - \hat{\mathcal{R}}_t^i \right]^2$$

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

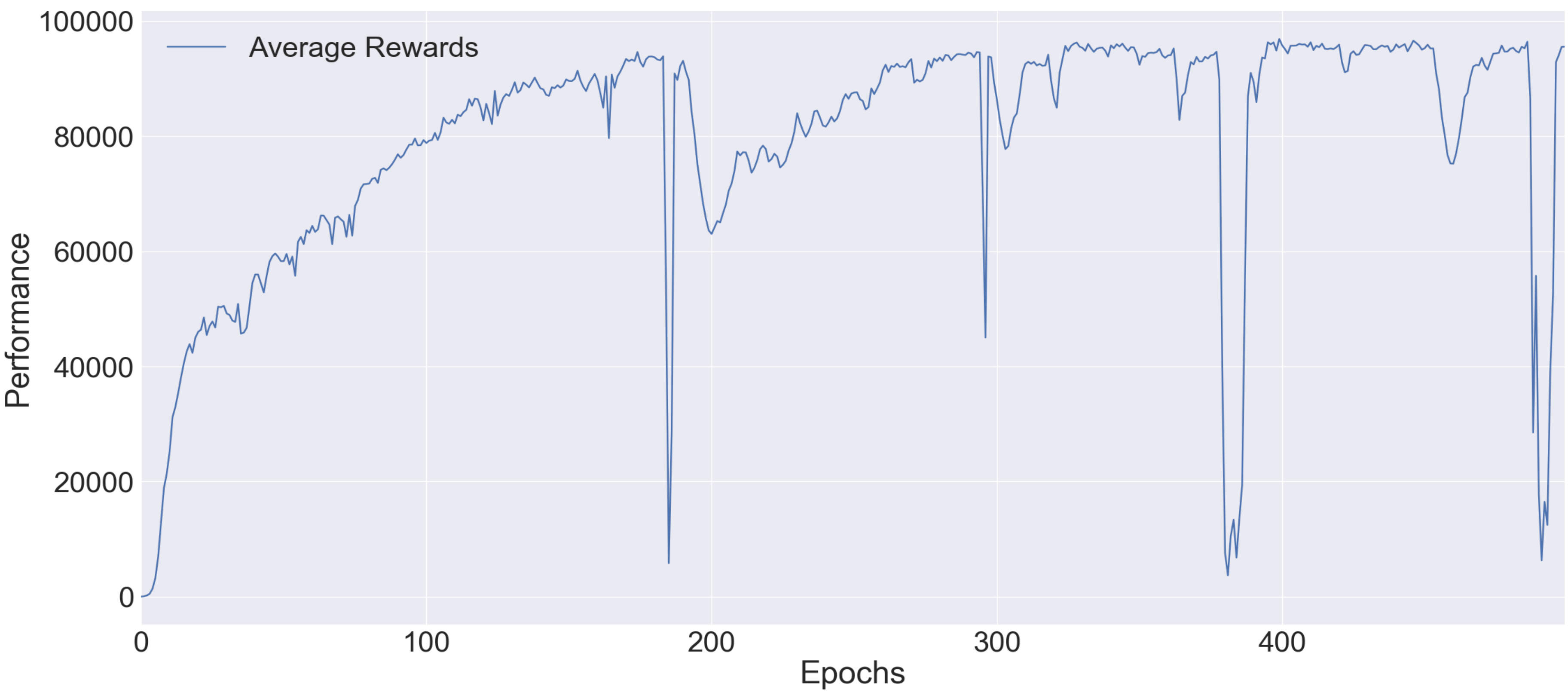
RL parameters

Parameters	Value
steps_per_epoch	26000
gamma	1
epochs	500
# episodes	125000
clip_ratio ϵ	0.2
pi_lr / vf_lr	$3e^{-4}$ / $1e^{-4}$
hidden layers	$[64, 64]$

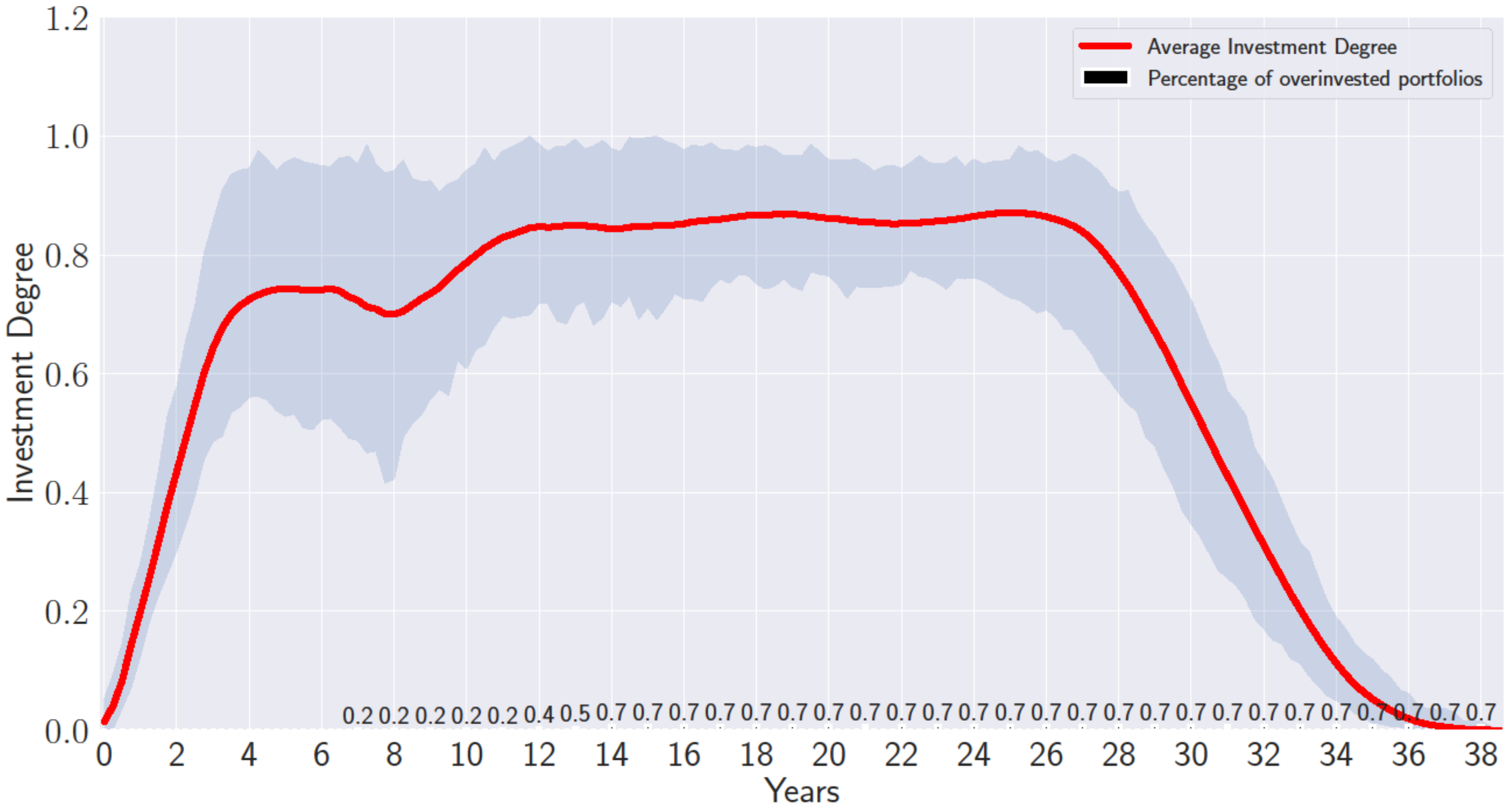
Simulation parameters

Parameters	Training	Validation
Cashflows frequency	quarterly	quarterly
Investment period	26 years	26 years
Funds per recommitment	4	4
Fund selection	ESG score	ESG score
Number of simulated portfolios (per evaluation)	250	1000
Distributed simulation	True	False

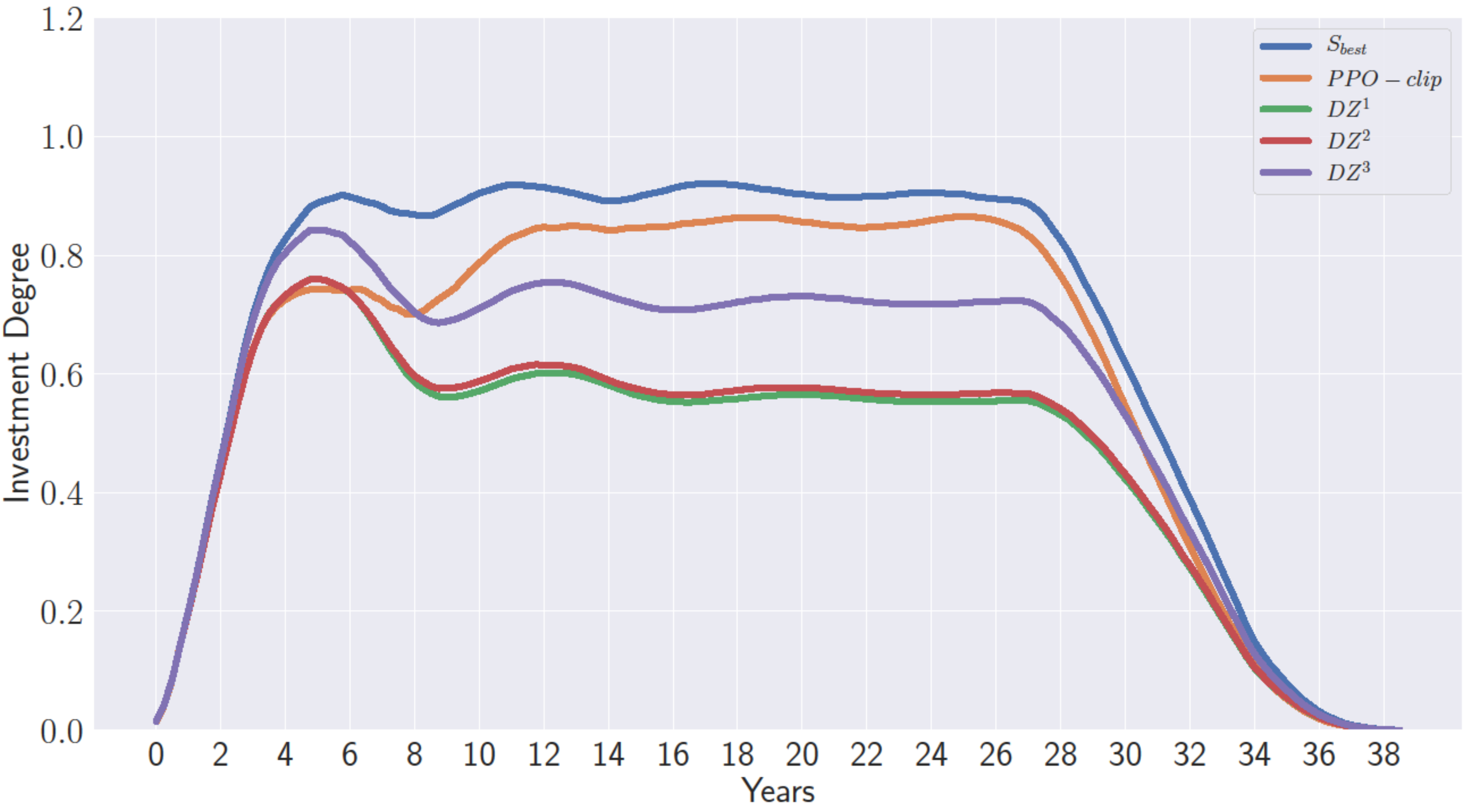
Rewards evolution



Best policy obtained with PPO-clip



Comparison with existing strategies



Conclusion

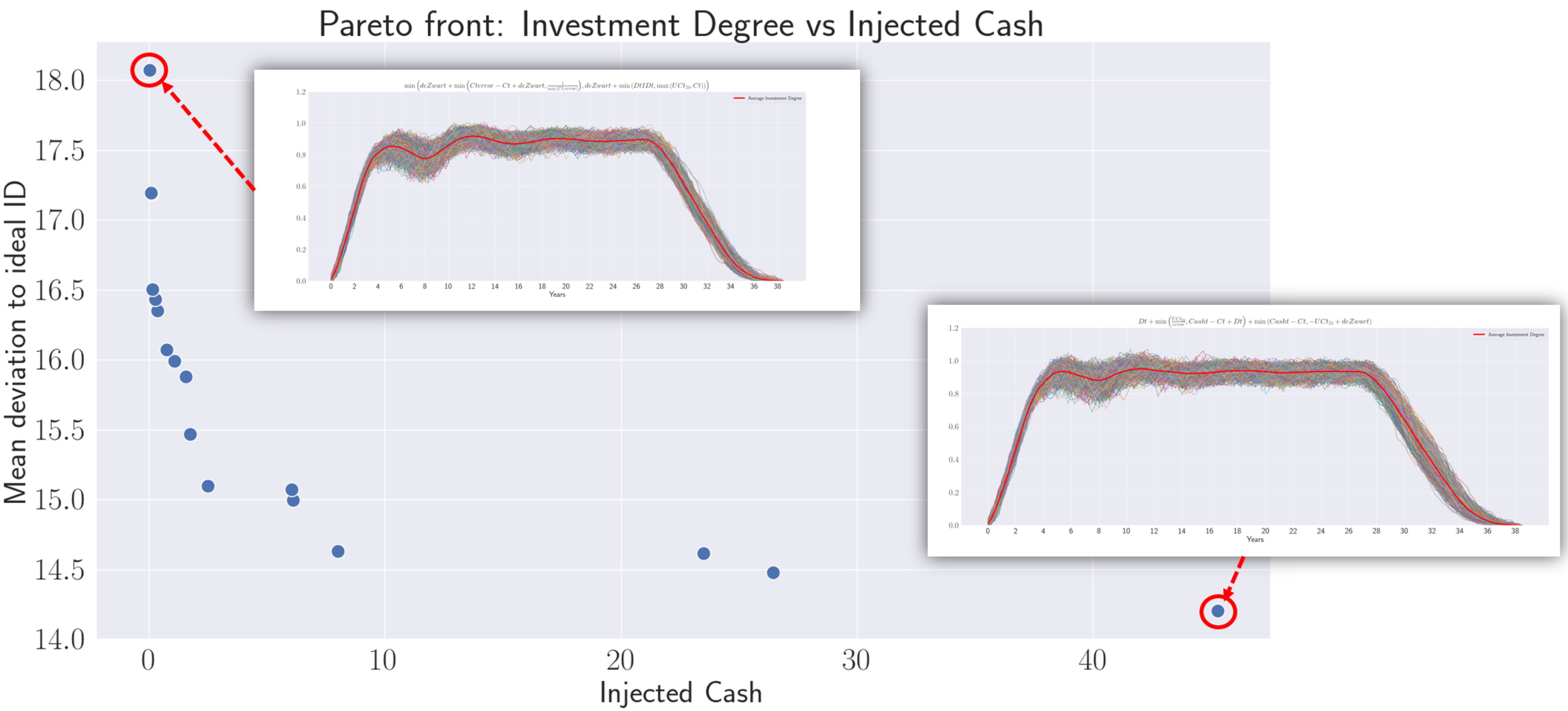
UNIVERSITÉ DU LUXEMBOURG

- 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