CVA Hedging with Reinforcement Learning

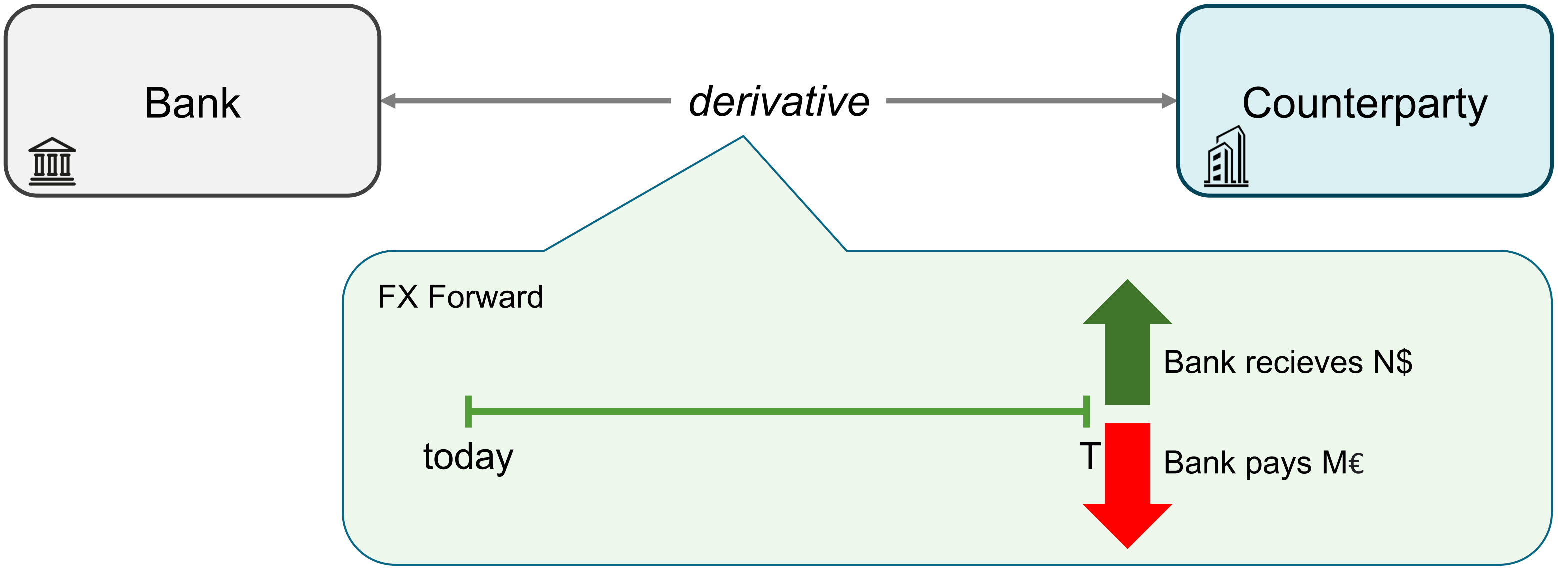
Edoardo Vittori

with Roberto Daluiso, Marco Pinciroli, Michele Trapletti
Intesa Sanpaolo SPA

ICAIF23
29 November 2023

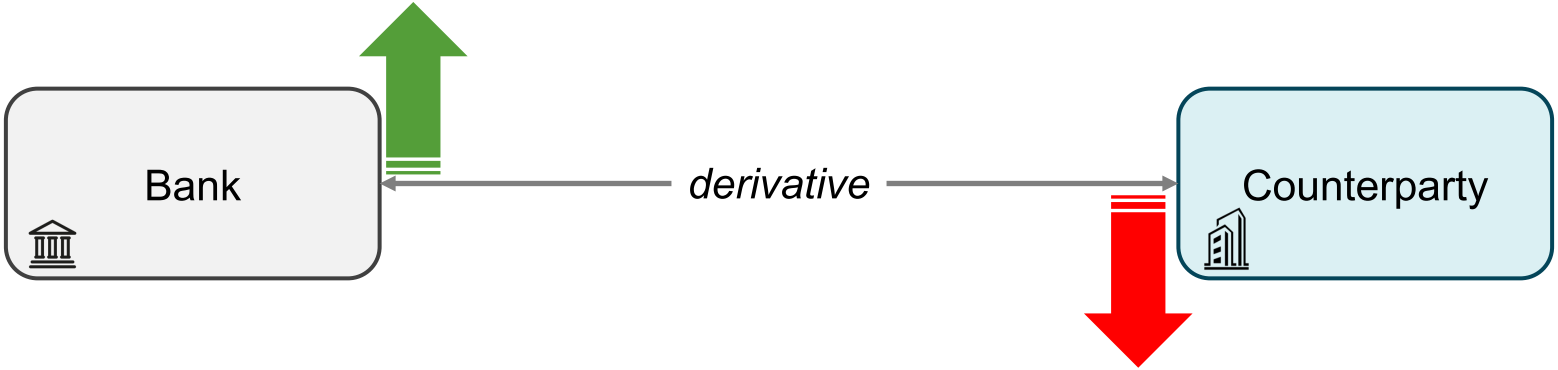
What is CVA?

We start with a derivative between a bank and a counterparty



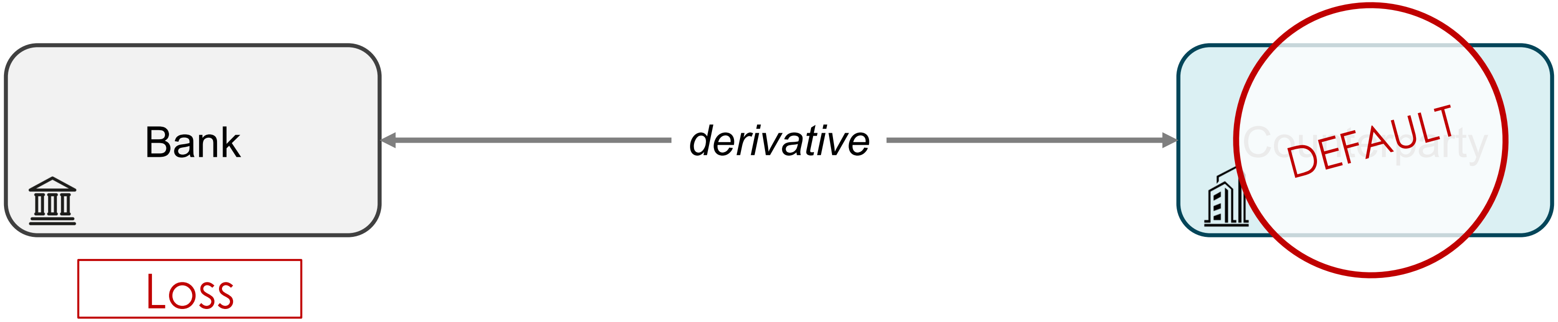
What is CVA?

Assume the derivative increases in value for the bank



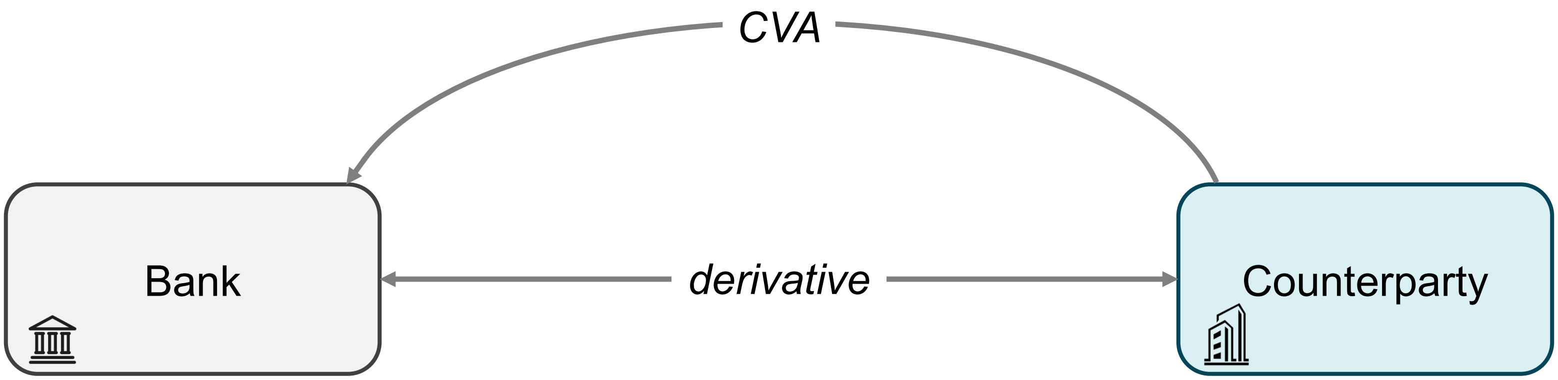
What is CVA?

The counterparty defaults and is not able to repay the bank



What is CVA?

CVA must be written on the balance sheet with mark to market evaluation

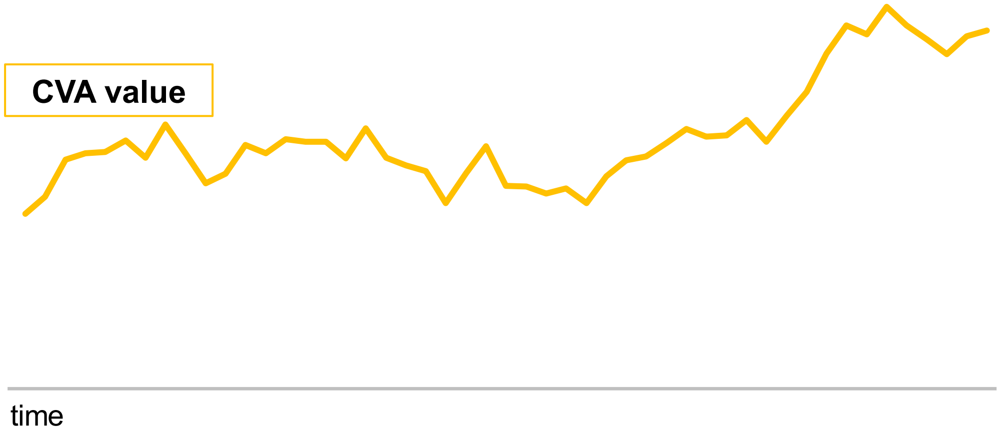


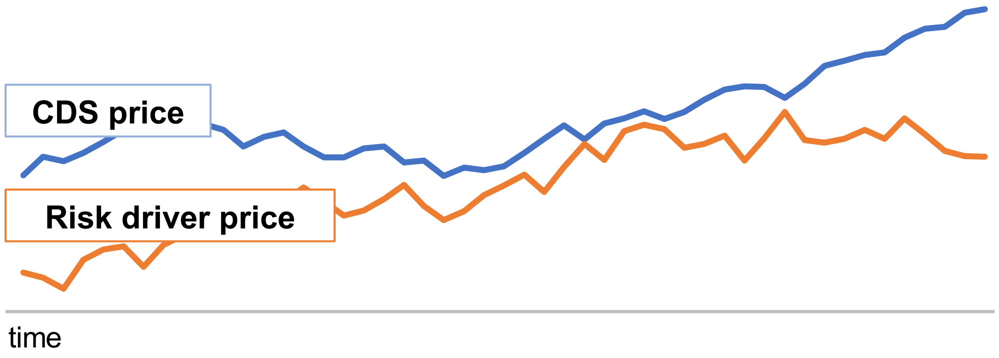
What is CVA

- Risk-neutral expectation of losses incurred if the counterparty of the derivative defaults
- $\mathrm{CVA} = \mathbb{E}[(\mathrm{LGD})\times(\text{Exposure at default})\times(\text{Probability of default})]$
- An option with multiple risk drivers and the possibility of default

Hedging CVA

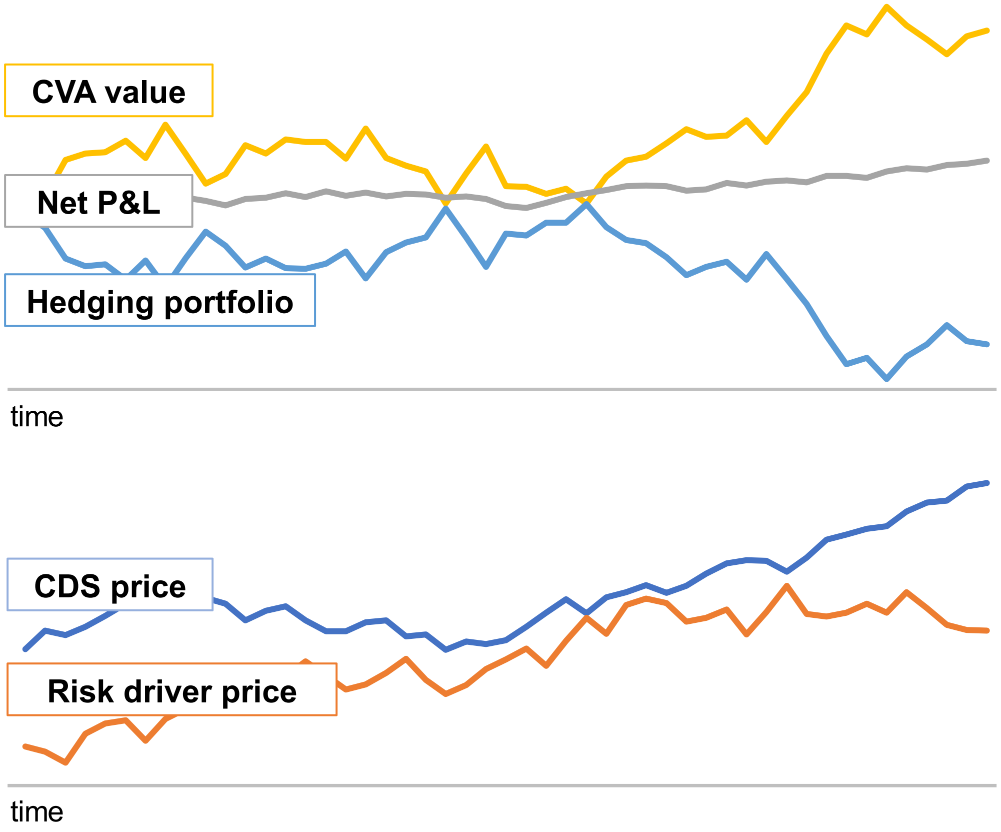
CVA changes during the lifetime of the derivative





Hedging CVA

Objective: achieve a low variance for CVA value + hedging portfolio



Hedging CVA definition

- Buy and sell the underlying risk drivers and CDS on the counterparty to offset the volatility of CVA

Challenges

- Costs when trading the hedging instruments
- Correlation between risk drivers
- Jump to default risk
- CVA pricing model misspecification

Reinforcement Learning for CVA Hedging

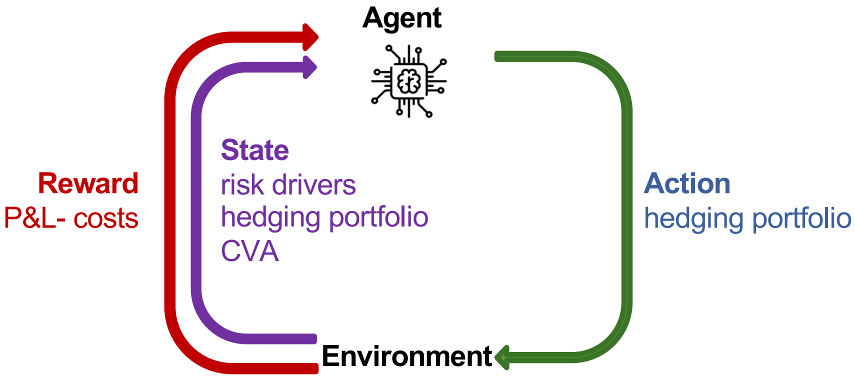
Reinforcement Learning introduction and MDP description

Reinforcement Learning Basics

- *MDP*: process which describes interaction between agent and environment
- *Objective*: find the policy π which maximizes the discounted sum of the rewards
- $\hat{J}_\pi = \mathbb{E}_\pi[\sum_i \gamma^i r_i]$

Hedging MDP

- *State*: risk drivers, time to CVA maturity, hedging portfolio, CVA
- *Action*: new hedging portfolio
- *Reward*: $\text{P\&L}_{cva}$ – $\text{P\&L}_{hdg}$ – transaction costs



Risk aversion in Reinforcement Learning

Return maximization and risk minimization tradeoff

Optimization objective

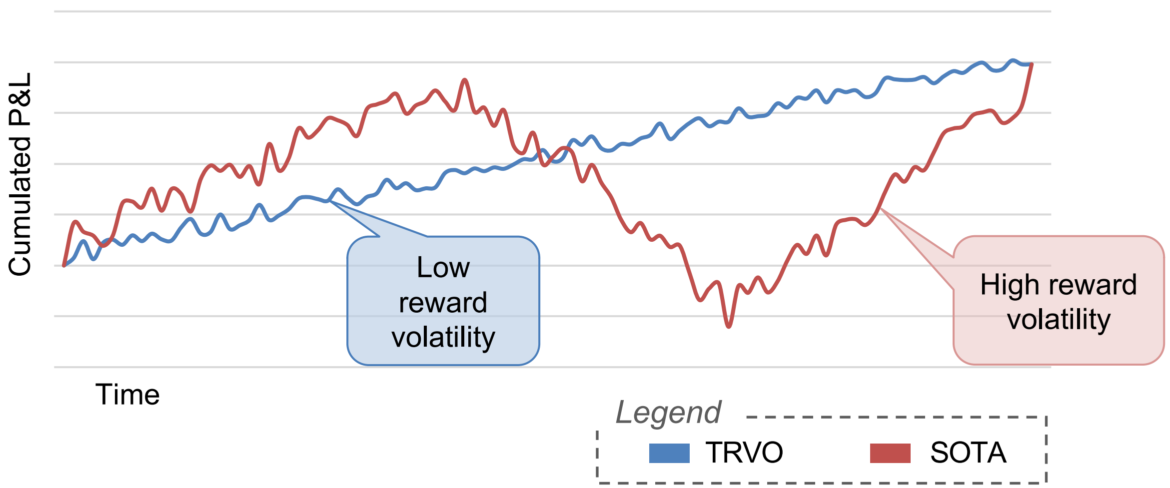
$$\hat{\eta}_\pi := \hat{J}_\pi - \beta \hat{\nu}_\pi^2$$

Reward volatility

$$\hat{\nu}_\pi^2 = \mathop{\mathrm{E}_\pi}_{s_0 \sim \mu}\left[\sum_{i=1}^{\varepsilon} \gamma^{i-1} (r_i - J_\pi)^2\right]$$

Return variance

$$\sigma_\pi^2 := \mathop{\mathrm{E}_\pi}_{s_0 \sim \mu}\left[\left(\mathcal{G} - \hat{J}_\pi\right)^2\right]$$



Experimental setting

We consider the CVA of an FX forward

Financial universe

- Derivative is an FX forward
- CDS with 5Y maturity
- Hedging instruments: CDS, FX spot EURUSD

RL setting

- 5 timesteps per day in trading hours (uneven time spacing)
- Horizon: earliest between default and 90 business days
- Batch size: 500
- Training iterations: 1500

Data generation

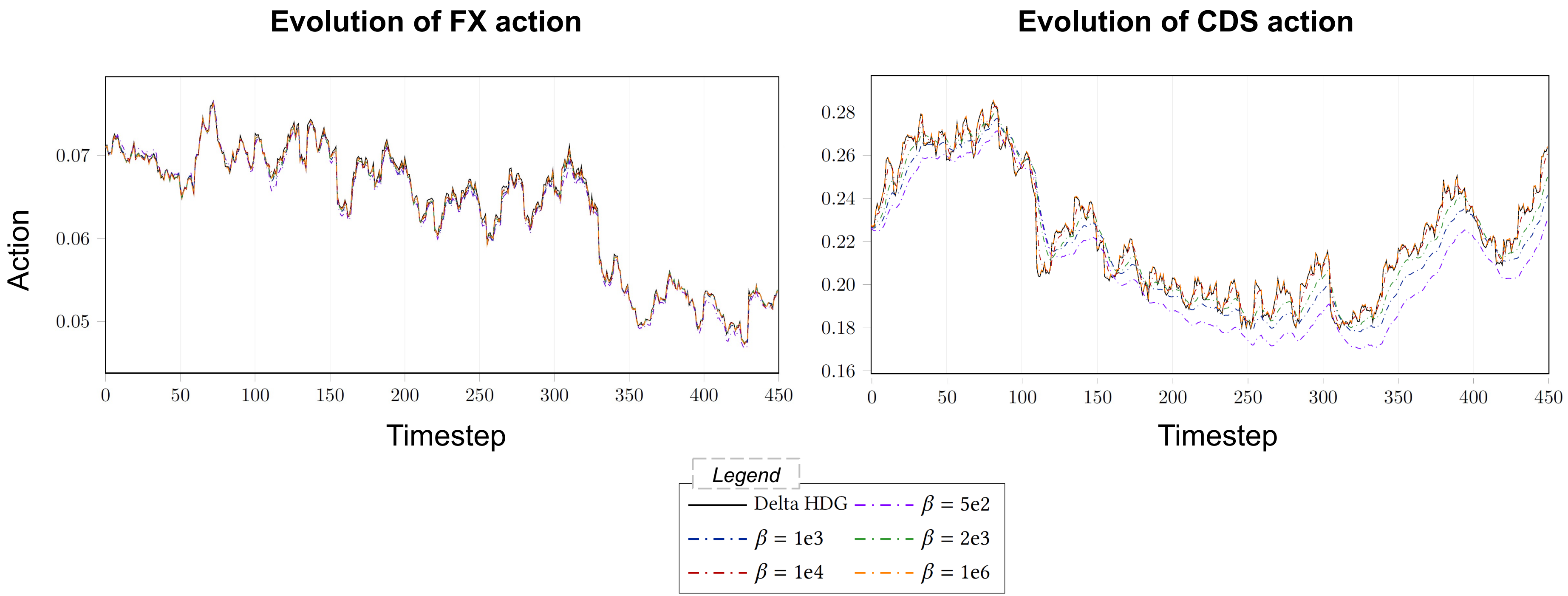
- FX: GBM model
- FX bid-ask: constant 1pip
- Credit: CIR intensity
- Credit bid-ask: ~ 10 to 60 bps depending on the experiment
- Correlation: 0 or 0.5 depending on the experiment

Baseline

- Delta hedge: policy which follows the first order derivatives of CVA w.r.t. the underlying risk drivers

Transaction costs and no correlation

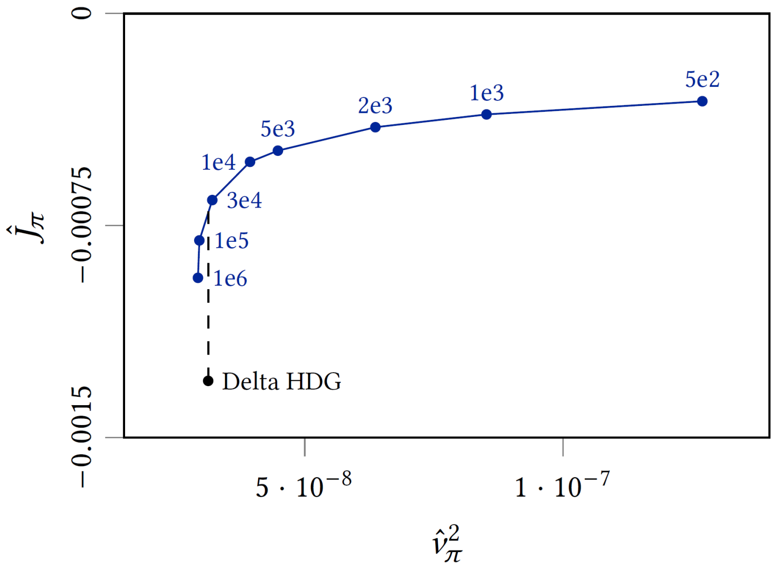
Experimental results – plot of policy



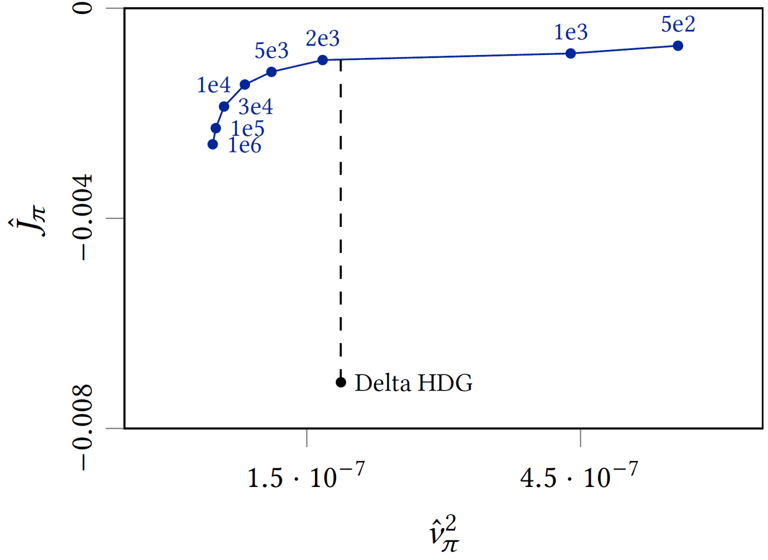
Transaction costs and no correlation

Experimental results – Pareto frontier

10 bps CDS bid-ask

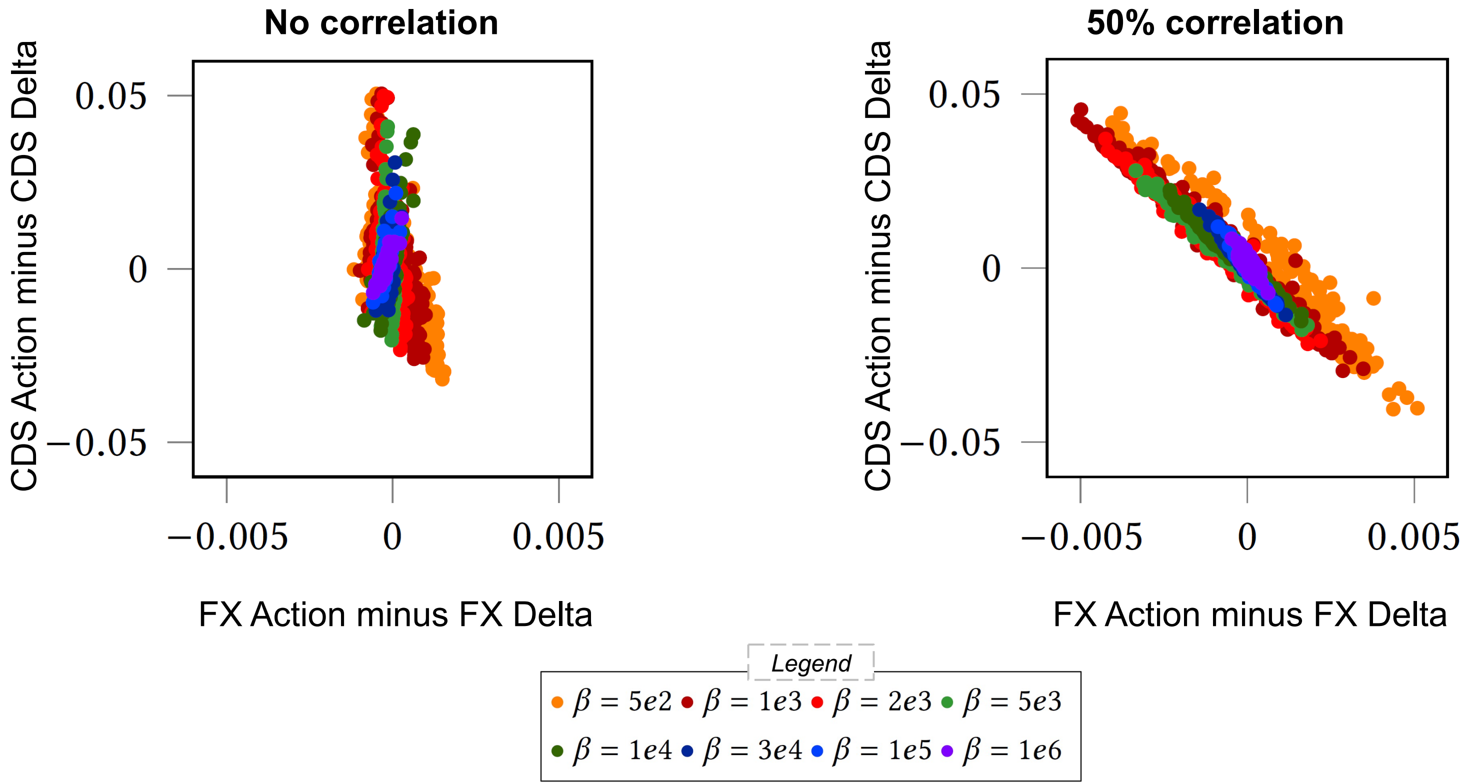


60 bps CDS bid-ask



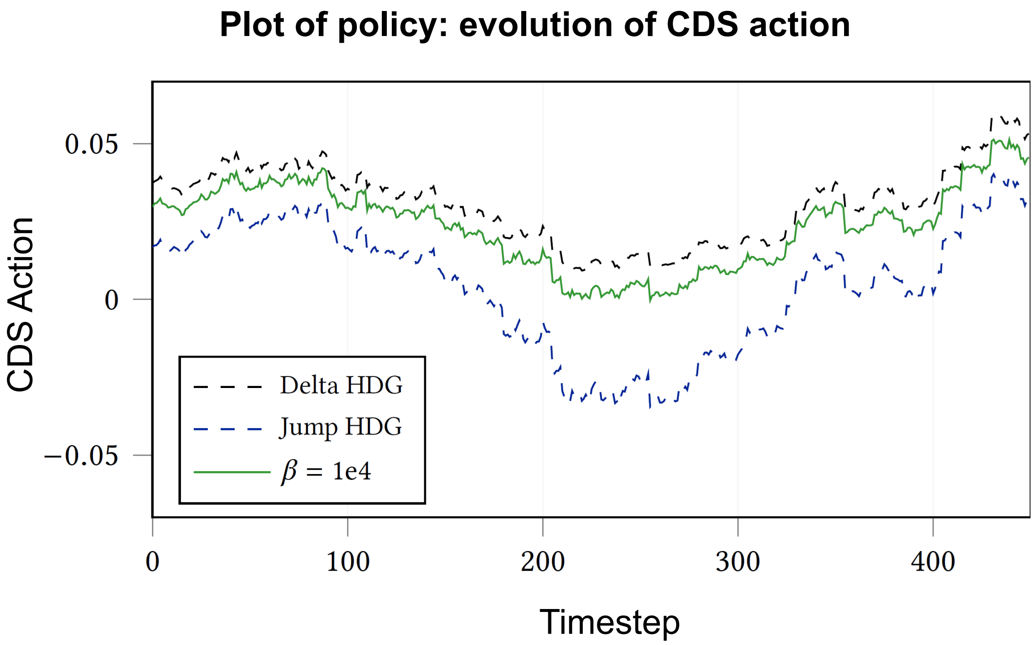
Transaction costs and correlation

Experimental results – comparison of agent actions with baseline actions



Defaults (no costs and no correlation)

Experimental results – plot of policy



Jump hedge

Baseline defined so that the notional of the hedging CDS perfectly offsets the P&L if the default is at the current timestep

CVA Hedging with Reinforcement Learning

We used risk averse reinforcement learning to hedge CVA

The optimized policies are superior to the standard delta hedge as they:

- reduce transaction fees
- exploit correlation
- generate an action between the delta hedge and the jump hedge when there are defaults

Edoardo Vittori
edoardo.vittori@intesasanpaolo.com