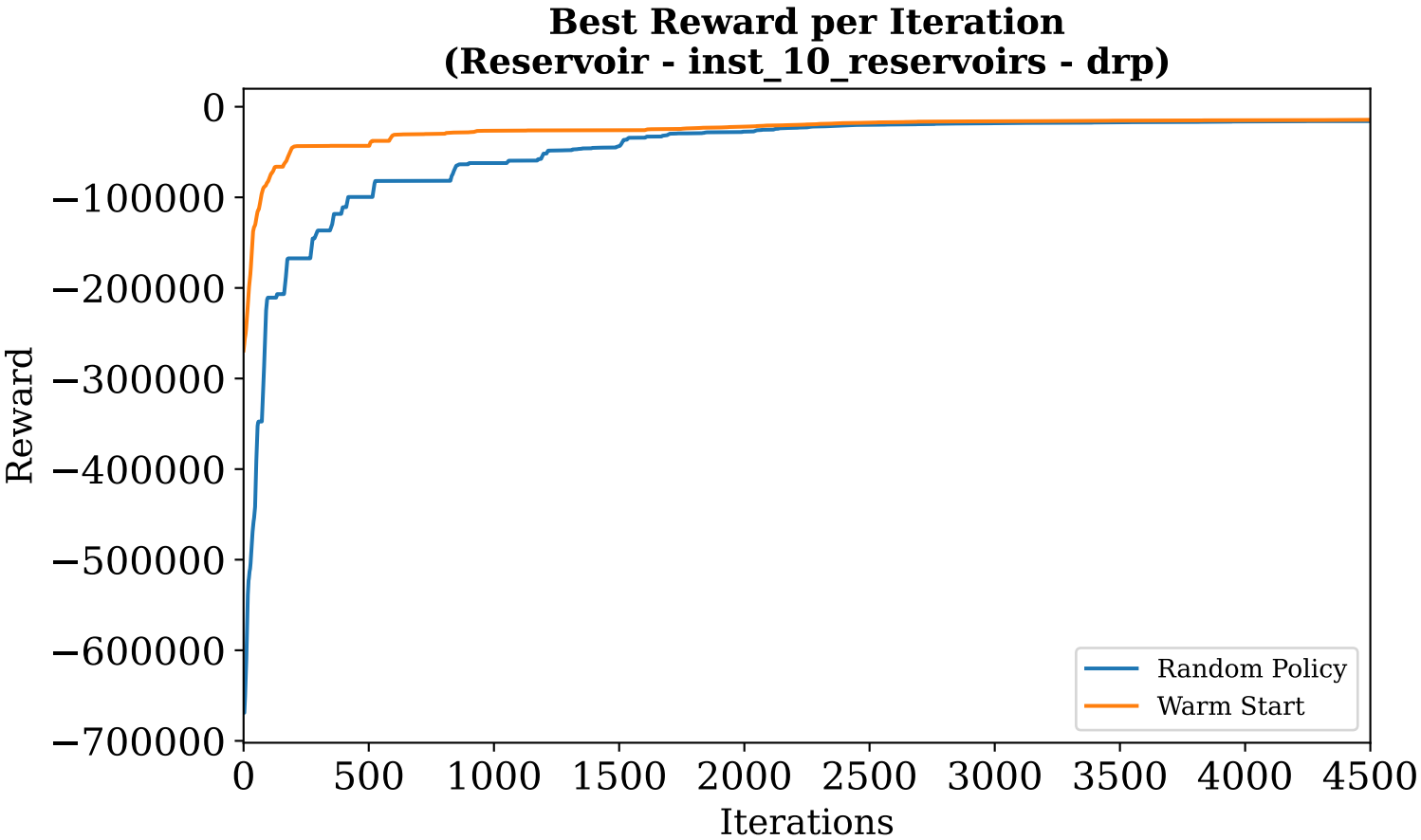
Best Reward per Iteration
(Reservoir - inst_10_reservoirs - drp)
Reward
Iterations
0
−100000
−200000
−300000
−400000
−500000
−600000
−700000
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Random Policy
Warm Start