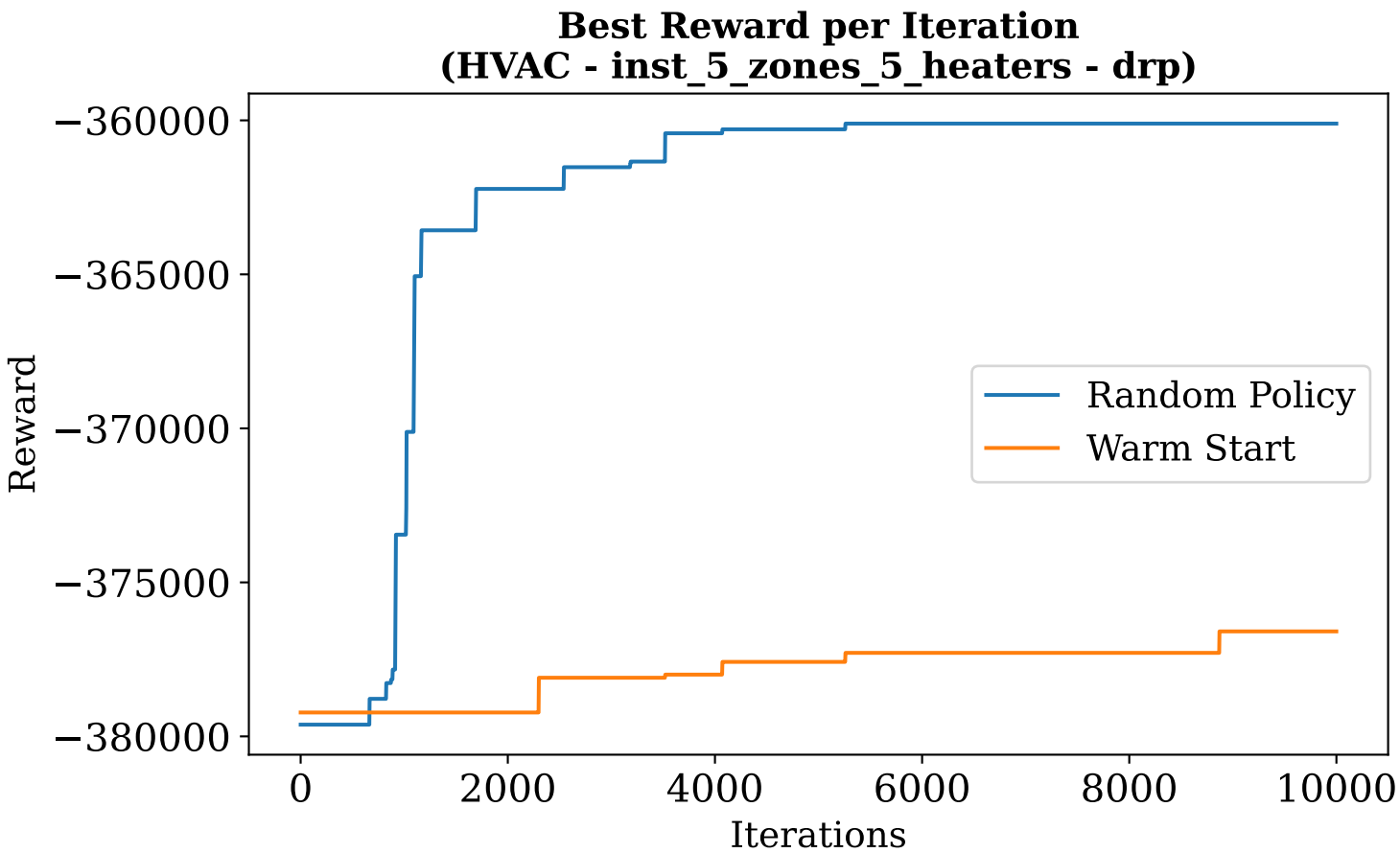
Best Reward per Iteration
(HVAC - inst_5_zones_5_heaters - drp)
−360000
−365000
−370000
−375000
−380000
Reward
0
2000
4000
6000
8000
10000
Iterations
Random Policy
Warm Start