erage reward per step for the 500 runs for the worst and best algo

average reward per step

greedy algorithm on non−stationary bandit
non−stationary algorithm with alpha = 0.5 and epsilon = 0.1

steps