Reward Trajectory for Uniform two UCB 0.60 -Algorithm Mean Reward DynamicEpsilonGreedy DynamicGreedy ThompsonSampling UCB1WithConstantOne UCB1WithConstantT 0.55 -0.50 -0 7500 2500 5000 10000 time