Reward Trajectory for Uniform cur 0.80 -0.75 -Algorithm 0.70 -Mean Reward DynamicEpsilonGreedy DynamicGreedy ExploreThenExploit ThompsonSampling UCB 0.65 -0.60 -Ó 2000 4000 6000 time