Reward Trajectory for Uniform TS, Dyn eps 0.8 -0.7 -Mean Reward Algorithm DynamicEpsilonGreedy ThompsonSampling 0.6 -0.5 -2000 5000 0 1000 3000 4000

time