## Number of rewards chosen by DBMR-BPI up to timestep T(N=20)400 300 200 平 鲁 皇 中 手 号 号 手 手 手 号 青 草 手 11 12 13 14 15 16 17 Reward id