

random MDP results; MSE of Q-function (left), and policy accuracy (right)

