

REINFORCE P-T vs. DQN P-T
 $\delta d = 0.2$ $\epsilon d = 0.3$

