

REINFORCE P-T vs. DQN P-T

$\delta d = 0.3$ $\epsilon d = 0.1$

