

REINFORCE P-T vs. DQN P-T

$\delta d = 0.4$ $\varepsilon d = 0.1$

