

REINFORCE P-T vs. DQN P-T

$\delta d = 0.2$ $\varepsilon d = 0.2$

