

RL non-batched experiments - q_a^π 's: near one, α : 0.001

