DQN C-T-V (σ=0.4)
action hist for p0 in]0.4, 0.2]