

$(\hat{Q}_h - Q^{\pi_{\text{target}}})^2$ MSE (10 episodes, $\gamma = 0.99$)

