Avg reward for a single move per trajectory, risk avers=0.9