Stag Hunt 5 Seed Episode Reward Mean Config Size Small 80 70 Policy Reward Mean 40 30 2 4 5 0 1e6 Total steps