IPD,4 and 5 players 4 player, 1 SHAPER 5 player, 1 SHAPER Reward Per Timestep Shaper 1 Shaper 1 8 6 NL<sub>1</sub> NL 1 6 NL<sub>2</sub> NL<sub>2</sub> NL3 NL<sub>3</sub> 4 NL4 2 2 0 250 500 750 1000 250 500 750 1000 1.0 1.0 State distribution 8.0 8.0 0.6 0.6 0.4 0.4 0.2 0.2 0.0 0.0 250 500 750 1000 250 500 750 1000 Eval episode Eval episode