IPD, 4 players 2 SHAPERs 3 SHAPERs Reward Per Timestep Shaper 1 Shaper 1 6 6 Shaper 2 Shaper 2 NL₁ Shaper 3 4 NL 2 NL₁ 2 0 0 250 500 750 1000 250 500 750 1000 1.0 1.0 -CD, CC DDD, C State distribution 8.0 8.0 DC, CC DD, CC 0.6 0.6 0.4 0.4 0.2 0.2 0.0 0.0 250 750 500 750 1000 250 500 1000 Eval episode Eval episode