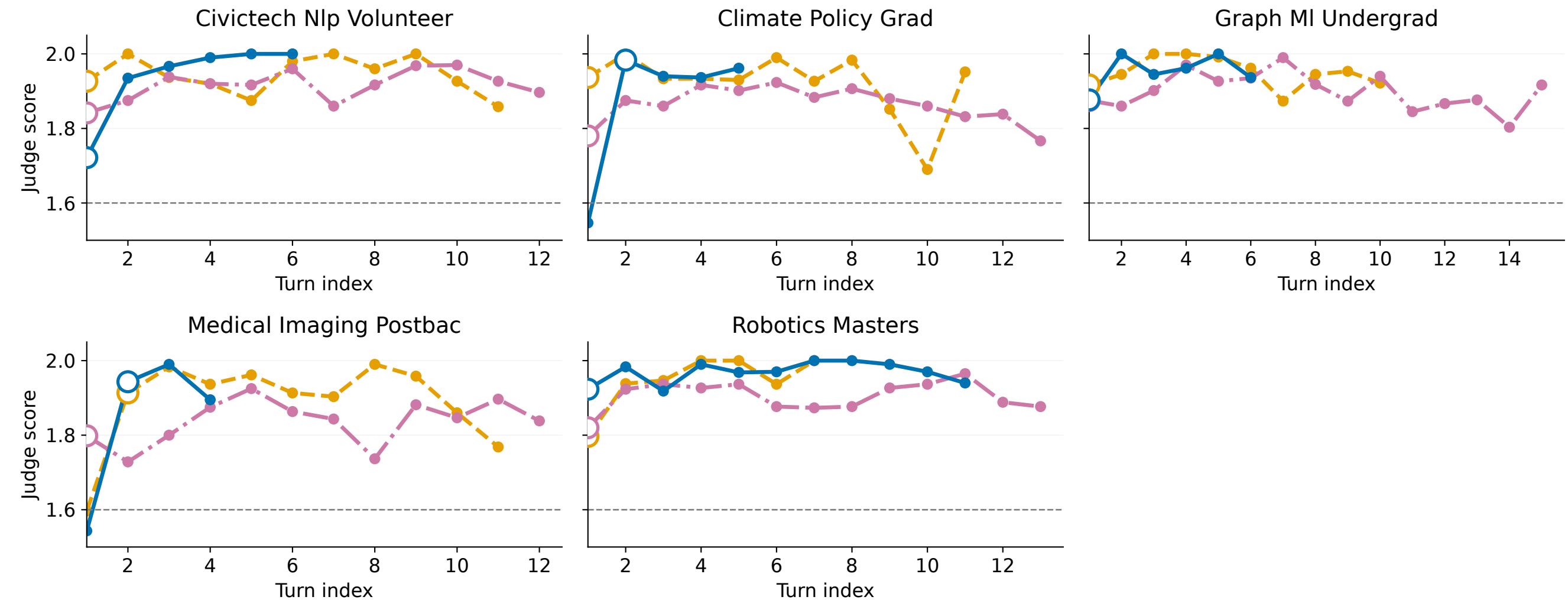


Per-scenario multi-turn trajectories



Hollow markers denote the first turn where score ≥ 1.6 .