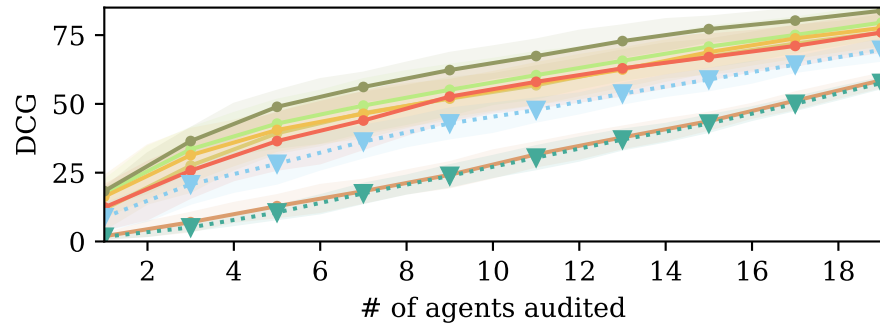
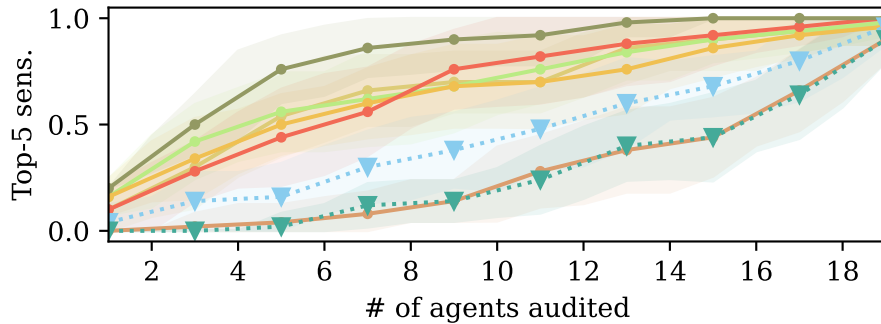
Ranking performance across auditing thresholds

Top-5 sensitivity (↑) — Top-5 sens. vs. # of agents audited

Discounted cumulative gain (DCG, ↑) — DCG vs. # of agents audited

Causal effect estimator: S-learner, DragonNet, T-learner, R-learner, S+IPW, PSM, Random, Payout-only