

Abstract

A principal incentivizes a group of agents to work by choosing a monitoring structure and a scheme of performance-contingent rewards. The monitoring structure partitions the set of agents into monitoring teams, each delivering a signal of joint performance. We show that unlike under Nash implementation, the principal always exhausts her monitoring capacity to optimally implement work as a unique outcome. Optimal monitoring teams are homogeneous between them: equally sized and with agents allocated in an anti-assortative fashion. Higher-effort-cost agents receive lower rents, and they tend to be monitored more closely than lower-effort-cost agents when the principal's allocation is constrained.