Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Economic Theory 147 (2012) 999–1034

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Combinatorial agency

Moshe Babaioff[a], Michal Feldman[b,*,1,2], Noam Nisan[c,3], Eyal Winter[d,4]

[a] *Microsoft Research – Silicon Valley, United States*

[b] *School of Business Administration and the Center for the Study of Rationality, The Hebrew University of Jerusalem, Israel*

[c] *School of Computer Science, The Hebrew University of Jerusalem, Israel*

[d] *Department of Economics, The Hebrew University of Jerusalem, Israel*

Received 23 September 2009; final version received 6 September 2011; accepted 6 November 2011

Available online 28 January 2012

Abstract

We study a combinatorial variant of the classical principal-agent model. In our setting a principal wishes to incentivize a team of strategic agents to exert costly effort on his behalf. Agents' actions are hidden and the principal observes only the outcome of the team, which depends stochastically on the complex combinations of the efforts by the agents. The principal seeks the mechanism that maximizes the principal's net revenue given an equilibrium behavior of the agents. We investigate the structure of the optimal mechanism for various production technologies as the principal's value from the project varies. In doing so we quantify the gap between the first-best and second-best solutions. Our results highlight the qualitative and quantitative differences between production technologies that exhibit complementarities and substitutabilities between the agents' actions. In comparing the first best with the second best we highlight the role of effort monitoring by the principal. As we shall see, the benefit from monitoring crucially depends on the underlying technology, with the two polar cases being perfect substitution and perfect complementarity.

* Corresponding author.

E-mail addresses: moshe@microsoft.com (M. Babaioff), mfeldman@huji.ac.il (M. Feldman), noam@cs.huji.ac.il (N. Nisan), mseyal@mscc.huji.ac.il (E. Winter).

1 Present address: School of Engineering and Applied Sciences (office 337), Harvard University, 33 Oxford Street, Cambridge, MA 02138, United States.

2 Supported in part by the Israel Science Foundation (grant number 1219/09), by the Leon Recanati Fund of the Jerusalem School of Business Administration, by the Google Inter-university center for Electronic Markets and Auctions and by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 274919.

3 Supported in part by a grant from the Israel Science Foundation and by the Google Inter-university center for Electronic Markets and Auctions.

4 Supported in part by the Thyssen Foundation and by the Google Inter-university center for Electronic Markets and Auctions.

0022-0531/$ – see front matter
doi:10.1016/j.jet.2012.01.010

JEL classification: C72; D82; D86

Keywords: Agency theory; Incentives; Contracts; Hidden-action; Moral hazard

1. Introduction

By "combinatorial agency" we refer to team environments where the stochastic success of a joint project is affected by the combination of effort decisions taken by the team members. A single principal, who is affected by the project's outcome but cannot observe the agents' effort decisions, contracts with the agents to maximize his profit. Team production and mechanisms for incentivizing teams are relevant for a variety of economic environments, as the functioning of organizations heavily relies on the optimal design of incentives for employees.

The purpose of this work is to study the relationship between the level of benefit the principal acquires from a successful project and the structure of his/her incentive contracts with the agents. As shall be seen later this relationship critically depends on the production technology (i.e., the mapping from agents' effort decisions to the probability of the project's success). The linkage between the principal's incentives and the design of agents' incentives is of considerable importance in changing environments. In commercial firms this can arise from changes in the prices of goods that are obtained from a successful completion of the project. In other instances this may be a result of changes in the subjective importance the principal attributes to the completion of the project. Some of our results will highlight the extent to which the optimal set of contracts are robust to such changes, and the way this robustness is linked to the underlying technology.

The problem of incentives in production teams was introduced by Holmstrom [4], who focused on social welfare maximization under budget balance constraints. In a follow-up work [1], a characterization for settings admitting compensation mechanisms satisfying budget balance, limited liability, and efficiency was provided. [11] considered the problem under sequential rather than simultaneous mechanisms.

Segal [7,8] studied properties of general multi-agent contracts. His analysis covers both the partial implementation case in which the optimal contract sustains the desirable outcome as a Nash equilibrium (not necessarily uniquely) as well as the case of full implementation (where the desirable outcome is sustained as a unique Nash equilibrium). Since these contracts are interpreted as trading contracts they do not involve hidden actions, and even within this framework Segal is addressing different questions than we do.

In departing from the above literature, we study the problem from the perspective of the principal seeking to maximize his own profit. This is also the approach taken by [2], who showed that in settings with repeated interactions, the principal can benefit from punishment strategies by peers. Apart from the different perspective, our framework departs from [4,1,11] in the set of research questions that we focus on. In particular, our purpose is to investigate the interdependency between the technology of team production and the structure of the optimal incentive mechanism. Specifically, we wish to study the relation between the production technology and the optimal set of agents that are employed under the optimal mechanism as well as the principal's value of information regarding agents' hidden actions.

Our model here generalizes Winter's model of team production [13]. In Winter's model a principal contracts with multiple agents. Agents make effort decisions that are mapped to a probability that the joint project ends successfully. We consider two frameworks for our analysis. In the case of hidden actions, the principal is informed only about the outcome of the entire project and not about individuals' effort choices. In the observable-actions case the principal is allowed to contract on agents' effort decisions. In [13], in both frameworks the principal attempts to induce all agents to exert effort at minimal cost. Our paper here builds on this model by taking a more general perspective and addressing a different set of questions. We assume that the principal generates a certain (monetary) reward from the success of the project. Depending on this reward he employs agents optimally and incentivizes them so as to maximize his expected net profit.

Our main interest in this paper is to study the influence of the organizational technology on two aspects of the optimal contract. The first concerns the question of how the technology affects the set of agents with whom the principal contracts, and the second concerns the value of information about the individuals' effort and the way it is affected by the technology. The latter issue can also be viewed as a benchmark for addressing the role of monitoring of agents' effort by the principal.

Our first objective in this framework is to study the way the principal's recruitment policy is affected by the technology of the project. Our analysis starts by comparing two polar cases of symmetric technologies, the AND technology and the OR technology. Both these technologies arise from individual tasks with the difference being that the success of the projects requires the successful completion of all the tasks under the AND technology, whereas it suffice that at least one task succeeds for the project to succeed under the OR technology. Clearly, in both technologies as the value of the project increases, so does the number of players that are recruited. However, it is less clear how the transition of the set of employed agents is affected by the technology. It turns out that the polarity of the two technologies is carried over to the transition: under the AND technology we have a phase transition which means that for a certain critical value for the project the optimal set of agents jumps from the empty set directly to the set of all agents. In contrast under the OR technology the transition is as smooth as it can be in covering all the $n + 1$ possible sizes (from zero to n) of optimal sets of agents.

In Section 4 we provide a full characterization of the technologies that breed a sharp phase transition (from the empty to the entire set of agents) in both the observable-actions framework as well as in the hidden-actions framework. Interestingly, in both frameworks the condition roughly requires that as we move from a contract with the entire set of agents to a contract with a strict subset the proportional expected cost of incentivizing the agents has a lower decline than that of the probability of success.

In addition to the single transition described above, we shall discuss the polar case of a large number of transitions, i.e., when the optimal set of agents that the principal contracts with grows smoothly with the principal's value of the project. Here, we focus on two special cases in which the number of transitions cannot exceed the number of agents, namely, anonymous technologies and any technology in the observable-actions case.[5] For these two special cases we obtain an elegant dual result to the case of single transition i.e., we give an exact characterization for the properties that ensure that the number of transition points equals the number of agents. For anonymous technologies in the observable-actions case the property corresponds to a notion of substitution (between agents' efforts).

[5] One can also consider general technologies in the hidden-actions case and ask for the maximal number of transition points that a technology can have. In principle, there may be up to $2^n - 1$ transitions, as there are 2^n different possible subsets. It was shown by [5] that the number can be very large, exponential in n, and [3] showed that even a non-anonymous OR technology can have an exponential number of transitions.

Our results point to some interesting connections between the type of returns to scale of the technology and the volatility of the set of agents recruited to the project, as a function of its value. In the observable-actions case increasing returns to scale (IRS) implies a single transition (from none to all). In the hidden-actions case IRS might not be sufficient to imply a single transition, however we can still assert that a strengthened version of IRS implies a single transition. We conclude that in both cases (observable and non-observable) a high level of complementarity in the technology implies that for some value of the project a small change will result with a drastic change in recruitment. This is in contrast to the case of substitution, as even for anonymous technologies decreasing returns to scale (DRS) implies maximal number of transitions (n).

In contrast to our model which assumes that agents act simultaneously, i.e., are not informed about their peers effort when they decide on their own, [12] considers a sequential model where agents have some (but not necessarily perfect) information about their peers. Interestingly, the properties of complementarity and substitution play an important role in analyzing the effect of such internal information on the cost of the optimal mechanism. Specifically, under complementarity more information among peers reduces the cost of the optimal mechanism, while under substitution it has no effect on this cost.

We end our analysis of the principal's optimal recruitment policy with a result that compares agents in terms of their effectiveness or importance in the production process. Two properties are defined here. The first is *productivity*: we say that agent i is more productive than agent j if i's marginal contribution exceeds that of j regardless of which group of agents they join in exerting effort. A dual relation is that of *productivity initiation*. We say that agent i is a better productivity initiator than j, if any other player, say k, contributes to a coalition containing i more than he contributes to the coalition obtained by replacing agent i with agent j. Interestingly, neither of these two relations among agents implies the other. We show that under hidden actions, if agent i is more productive than j and also a better production initiator than j, then regardless of the project's value for the principal, j will be a member of the selected team only if i is a member. We argue that this property applies only if i "dominates" j in both relations: just one of them will not suffice.

Our second topic of investigation concerns the principal's loss of profit due to the lack of information about agents' effort decisions. Since these decisions are hidden, they are uncontractable, which forces the principal to implement a second-best outcome. We term this loss as the *price of unobservability* (*POU*), which for a given technology is defined as the worst ratio (over all possible values of the project) between the principal's net profit under perfect observability of effort and under hidden efforts.

It turns out that the analysis of this issue strongly builds on our analysis of the principal's optimal recruitment policy. We find that the price of unobservability is unbounded for the class of project technologies involving increasing returns to scale (i.e., those having complementarities among agents). This result contrasts with the one holding for the case of technologies involving decreasing returns to scale (i.e., technologies having substitutability among agents), where the price of unobservability is always bounded by a linear function of the number of agents. The distinction between complementarity and substitution here has an interesting practical implication. If the principal can introduce costly monitoring into the organization in order to monitor agents' efforts, he may find such monitoring particularly effective under complementarity, where the loss in profit due to hidden actions can be arbitrarily close to 100%.

Properties of optimal contracts are important from both a descriptive perspective as well as a normative one. First, they allow us to examine, through empirical results, whether specific institutions are operating under optimal mechanisms without having to specify the optimal mechanism explicitly or even without fully specifying the technology. From the normative perspective they allow us to derive rules or amendments that if introduced may shift the outcome towards efficiency in spite of the mechanism still remaining non-optimal.

In Section 2 we present the formal model and the definitions of the main concepts. Section 3 provides the analysis of the two polar technologies of AND and OR and demonstrates the relation between the transitions of the optimal set of contracts and the properties of substitution and complementarity. In Section 4 we analyze the general framework and provide results regarding the transition of the contracts as the value of the project varies. Section 5 discusses the role of productivity and productivity initiations in the design of the optimal contracts. Section 6 provides the analysis of the price of unobservability. Section 7 concludes.

2. Model and preliminaries

We begin by presenting our model and some preliminaries. All missing proofs from this section appear in Appendix A.

2.1. The model

A principal employs a set of agents N of size n. Each agent $i \in N$ has a possible set of actions A_i, and a cost (effort) $c_i(a_i) \geqslant 0$ for each possible action $a_i \in A_i$ ($c_i : A_i \to \Re_+$). In this paper we consider the case where for each agent i the action space is binary: every agent chooses between action 0 (low effort) and 1 (high effort). Normalizing the cost of the low effort action to 0, the cost function of agent i can be represented by a scalar $c_i > 0$ denoting the cost of exerting high effort. The vector of costs is denoted by $\vec{c} = (c_1, c_2, \dots, c_n)$.

Let O denote the set of possible outcomes. We consider the case where the outcome space has only two states (binary outcome): 0 (project failure) and 1 (project success). The actions of all agents determine, in a probabilistic way, an outcome $o \in O$, according to a *success function* $t : A_1 \times \dots \times A_n \to \Delta(O)$ (where $\Delta(O)$ denotes the set of probability distributions on O). Since each agent's action space as well as the outcome space are binary, we observe that the success function t can be looked at as a function of the form $t : \{0,1\}^n \to [0,1]$, where $t(a_1, \dots, a_n)$ denotes the probability of project success where agent i with $a_i = 0$ does not exert effort (and incurs no cost), and agent i with $a_i = 1$ does exert effort (and incurs a cost of c_i).

A production *technology* is a pair, $(t, \vec{c})$, of a success function, t, and a vector of costs, $\vec{c} = (c_1, c_2, \dots, c_n)$. The principal has a certain value for each possible outcome, given by the value function $v : O \to \Re$. Normalizing the value of project failure to 0, the value function can be represented by a scalar $v > 0$ which denotes the principal's value for a successful project. As we will consider only risk-neutral agents in this paper,[6] we will also treat v as a function on $\Delta(O)$, by simply considering the expected value given the outcome distribution.

In the hidden-actions case, actions of the agents are not observed by the principal, but the final outcome o is visible to him and to others (in particular the court), and he may design enforceable contracts based on the final outcome. Thus the contract for agent i is a function (payment) $p_i : O \to \Re$; again, we will also view p_i as a function on $\Delta(O)$. We assume that the principal can pay the agents but not fine them (known as the *limited liability* constraint).

[6] The first-best solution is not achievable here despite the fact that we assume risk-neutral agents since we impose the limited-liability constraint. The risk-averse case would obviously be a natural second step in the research of this model.

Normalizing the payment to agent i when the project fails to 0, we can now represent a contract with agent i by a scalar value $p_i \geqslant 0$ that denotes the payment that i gets in case of project success. When the lowest-cost action has zero cost (as we assume), limited liability immediately implies that the participation constraint (also known as the individual rationality constraint) holds.[7]

Given this setting, the agents have been put in a game where the utility of agent i under the vector of actions $a = (a_1, \dots, a_n)$ is given by $u_i(a) = p_i \cdot t(a) - c_i \cdot a_i$.[8] The principal's problem (which is our problem in this paper) is how to design the contracts p_i so as to maximize his own expected utility $u(a) = t(a) \cdot (v - \sum_{i \in N} p_i)$, where the actions $a_1, \dots, a_n$ are at a Nash equilibrium.[9] In the case of multiple Nash equilibria, in our model we let the principal choose the desired one and "suggest" it to the agents, thereby focusing on the "best" Nash equilibrium.[10] Finally, the *social welfare* for $a \in A$ is the sum of the principal's utility and the agents' utilities, given by $u(a) + \sum_{i \in N} u_i(a) = v \cdot t(a) - \sum_{i \in N} c_i \cdot a_i$.

As we wish to concentrate on motivating agents, rather than on the coordination between them, we assume that more effort by an agent always leads to a better probability of success; i.e., the success function t is strictly monotone. Formally, if we denote by $a_{-i} \in A_{-i}$ the $(n-1)$-dimensional vector of the actions of all agents excluding agent i, i.e., $a_{-i} = (a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$, then a success function must satisfy

$$\forall i \in N,\ \forall a_{-i} \in A_{-i} \quad t(1, a_{-i}) > t(0, a_{-i}).$$

Additionally, we assume that $t(a) > 0$ for any $a \in A$ (or equivalently, $t(0, 0, \dots, 0) > 0$).

We next define the marginal contribution of an agent given the actions of the others; that notion is important in characterizing the optimal contract and in defining different properties of technologies.

Definition 2.1. The marginal contribution of agent i, given $a_{-i} \in A_{-i}$, is

$$\Delta_i(a_{-i}) = t(1, a_{-i}) - t(0, a_{-i}).$$

$\Delta_i(a_{-i})$ is the increase in success probability due to agent i changing from not exerting effort to exerting effort, given the effort levels of the other agents. Note that since t is monotone, $\Delta_i(a_{-i})$ is always positive.

Given two vectors $a, b \in \{0,1\}^n$, we denote $a_{-i} < b_{-i}$ if for every agent $j \neq i$ it holds that $a_j \leqslant b_j$, and for some $j \neq i$ it holds that $a_j < b_j$. Some of our results concern technology success functions that satisfy one of the following standard properties.

---

7 In the observable-actions case the situation is much simpler: as actions can be monitored the principal can pay an agent only if the agent exerts effort, and paying the cost of the action is sufficient to ensure incentive computability and individual rationality.

8 One could think of a different model in which the agents have intrinsic utility from the outcome and payments may not be needed, as in [10,9].

9 This is the main difference between our work and [13], who assumes that the principal always wishes to induce an equilibrium in which all the agents exert their maximal effort.

10 In this paper we focus on the "best" Nash equilibrium. One may alternatively require the existence of a *unique* equilibrium as in [13], or alternatively, attempt to model some kind of an extensive game between the agents, as in [6,10,9].

Definition 2.2. A technology success function[11] t exhibits

- (strictly) **increasing returns to scale (IRS)**[12] if for every i and every $a_{-i} < b_{-i}$ it holds that $\Delta_i(a_{-i}) < \Delta_i(b_{-i})$;
- (strictly) **decreasing returns to scale (DRS)**[13] if for every i and every $a_{-i} < b_{-i}$ it holds that $\Delta_i(a_{-i}) > \Delta_i(b_{-i})$.

Intuitively, IRS means that more effort by the other agents increases the influence of an agent on the success of the project. This is a form of *complementarity* among the agents. On the other hand, DRS means that more effort by the other agents decreases the influence of an agent on the success of the project. This is a form of *substitutability* among the agents.

2.2. The optimal contract

We next discuss the structure of the optimal contract that maximizes the principal's utility in the hidden-actions case.

We start with some simple observations. The best action, $a_i \in A_i$, of agent i can be easily determined as a function of what the others do, $a_{-i} \in A_{-i}$, and his contract p_i.

Claim 2.3. *Given a profile of actions $a_{-i} \in A_{-i}$, agent i's best strategy is $a_i = 1$ if $p_i \geqslant \frac{c_i}{\Delta_i(a_{-i})}$, and is $a_i = 0$ if $p_i \leqslant \frac{c_i}{\Delta_i(a_{-i})}$. (In the case of equality the agent is indifferent between the two alternatives.)*

As $p_i \geqslant \frac{c_i}{\Delta_i(a_{-i})}$ if and only if $u_i(1, a_{-i}) = p_i \cdot t(1, a_{-i}) - c_i \geqslant p_i \cdot t(0, a_{-i}) = u_i(0, a_{-i})$, i's best strategy is to choose $a_i = 1$ in this case.

This allows us to specify the contracts that are the principal's optimal, for inducing a given equilibrium.

Observation 2.4. *The best contracts (for the principal) that induce $a \in A$ as an equilibrium are $p_i = 0$ for agent i who exerts no effort ($a_i = 0$), and $p_i = \frac{c_i}{\Delta_i(a_{-i})}$ for agent i who exerts effort ($a_i = 1$).*

In this case, the expected utility of agent i who exerts effort is $c_i \cdot \frac{t(0,a_{-i})}{\Delta_i(a_{-i})}$, and 0 for an agent who shirks. The principal's expected utility is given by $u(a, v) = (v - P) \cdot t(a)$, where P is the total payment in case of success, given by $P = \sum_{i|a_i=1} \frac{c_i}{\Delta_i(a_{-i})}$.

Note that as we assume that $t(a) > 0$ for any $a \in A$, an agent that exerts effort has positive utility, and the principal does not get all the social welfare. We say that the principal *contracts with* agent i if $p_i > 0$ (and $a_i = 1$ in the equilibrium $a \in A$). The principal's goal is to maximize his utility given his value v, i.e., to determine the profile of actions $a^* \in A$, which gives the highest value of $u(a, v)$ in equilibrium. Choosing $a \in A$ corresponds to choosing a set S of agents that exert effort ($S = \{i \mid a_i = 1\}$). We call the set of agents $S^*(v)$ that the principal contracts with

[11] If the success function of the technology with identical costs (t, c) exhibits some property, we also say that the technology exhibits the same property.

[12] A success function t exhibits IRS if and only if it is a super-modular function.

[13] A success function t exhibits DRS if and only if it is a sub-modular function.

in a^* ($S^*(v) = \{i \mid a_i^* = 1\}$) an *optimal contract* for the principal at value v. We sometimes abuse notation and denote $t(S)$ instead of $t(a)$, when S is exactly the set of agents that exert effort in $a \in A$.

We next present simple yet fundamental observations about changes to the optimal contracts caused by an increase in the principal's value. We observe that such an increase cannot result in a decrease in the principal's expected utility, in the success probability of the project, or in the total payment to the agents.

Lemma 2.5 *(Monotonicity lemma). For any technology $(t, \vec{c})$, in both the hidden-actions and the observable-actions cases, all the following are monotonically non-decreasing with the value*: (i) *the expected utility of the principal at the optimal contracts,* (ii) *the success probability of the optimal contracts, and* (iii) *the expected payment of the optimal contracts.*

In both the hidden-actions and the observable-actions cases, for every contract with a set T the utility of the principal is linear in v, with a slope of $t(T)$. The lemma implies that the utility of the principal at the optimal contract is a convex piecewise-linear function of v. A *transition point* v is a value in which this function changes its slope (an indifference point between two optimal contracts such that one contract is not optimal for higher values, and the other contract is not optimal for lower values).

2.3. Price of unobservability

A natural yardstick by which to measure this decision is the observable-actions case, i.e., when the principal observes the agents' individual actions, and thus can make the payment contingent on their actions. In this case, the principal motivates an agent to exert effort by paying the agent his cost. Therefore, the principal's utility under the profile $a \in A$ is $t(a) \cdot v - \sum_{i|a_i=1} c_i$. Recall that the social welfare under a profile a is also $t(a) \cdot v - \sum_{i|a_i=1} c_i$. Thus, in the observable-actions case, the principal's utility is equal to the social welfare, and the principal will simply choose the profile that optimizes the global efficiency ("first best").[14]

The worst ratio between the social welfare in this observable-actions case and the social welfare for the profile $a \in A$ chosen by the principal in the hidden-actions case, may be termed the *social price of unobservability*. To define this notion formally we first establish some notations.

Given a technology $(t, \vec{c})$, recall that $S^*(v)$ denotes an optimal contract in the hidden-actions case. Let $S^*_{oa}(v)$ denote an optimal contract in the observable-actions case, when the principal's value is v. An optimal contract in the observable-actions case is a contract that maximizes the principal's utility in the observable-actions case; note that such a contract also maximizes the social welfare.

Definition 2.6. The *social price of unobservability $POU_S(t, \vec{c})$* of a technology $(t, \vec{c})$ is defined as the worst ratio (over v) between the total social welfare in the observable-actions case and the hidden-actions case:

[14] Note that the first best can also be achieved if we relax the assumption of full observability by assuming instead that agent i's effort decision is observed/monitored with probability $q_i > 0$. In this case agent i is paid c_i/q_i whenever he is monitored exerting effort and zero otherwise. This yields an expected payoff of c_i to agent i who exerts effort, and not more.

$$POU_S(t, \vec{c}) = Sup_{v>0} \frac{t(S^*_{oa}(v)) \cdot v - \sum_{i \in S^*_{oa}(v)} c_i}{t(S^*(v)) \cdot v - \sum_{i \in S^*(v)} c_i},$$

where $S^*_{oa}(v)$ is any optimal contract in the observable-actions case, and $S^*(v)$ is an optimal in the hidden-actions case that has the lowest social welfare (worst one).

Similarly, the worst ratio between the *principal's utility* in the observable-actions case and in the hidden-actions case may be termed the *principal's price of unobservability*.

Definition 2.7. The *principal's price of unobservability* $POU_P(t, \vec{c})$ of a technology $(t, \vec{c})$ is defined as the worst ratio (over v) between the principal's utility in the observable-actions case and the hidden-actions case:

$$POU_P(t, \vec{c}) = Sup_{v>0} \frac{t(S^*_{oa}(v)) \cdot v - \sum_{i \in S^*_{oa}(v)} c_i}{t(S^*(v))(v - \sum_{i \in S^*(v)} \frac{c_i}{t(S^*(v)) - t(S^*(v) \setminus \{i\})})},$$

where $S^*_{oa}(v)$ is any optimal contract in the observable-actions case, and $S^*(v)$ is any optimal contract in the hidden-actions case.

When the technology $(t, \vec{c})$ is clear in the context we will use POU_S and POU_P respectively to denote the social price of unobservability and the principal's price of unobservability for technology $(t, \vec{c})$. Note that POU_S and POU_P are at least 1 for any technology. In addition, the principal's price of unobservability is always (weakly) larger than the social price of unobservability, as the following observation indicates.

Observation 2.8. *For any technology $(t, \vec{c})$ it holds that $POU_P \geqslant POU_S$.*

The definitions of POU_P and POU_S seem to require finding the supremum over a continuum of values. We next show that this is not the case and that one needs only to find the maximum over finitely many candidate points. These points are points of transitions between optimal contracts (there are only finitely many such points as each contract can only be optimal on an interval by Lemma 2.5).

Lemma 2.9. *For any given technology $(t, \vec{c})$ the social price of unobservability is obtained at some value v which is a transition point of either the hidden-actions or the observable-actions case. The same holds for the principal's price of unobservability.*

2.4. Natural restrictions

As we would like to focus on results that are derived from properties of the success function, in the sequel (unless otherwise stated) we will restrict our attention to the case where all agents have an *identical cost* c, that is, $c_i = c$ for all $i \in N$. We denote a technology $(t, \vec{c})$ with identical costs by (t, c).

An additional natural restriction for which we provide some results is on technologies that are anonymous. A success function t is called *anonymous* if it is symmetric with respect to the agents. I.e., $t(a_1, \ldots, a_n)$ depends only on $\sum_{i \in N} a_i$ (the number of agents that exert effort). A technology (t, c) is *anonymous* if t is anonymous and the cost c is identical for all agents

($\exists c$ s.t. $c_i = c$, $\forall i$). As for an anonymous success function t only the number of agents that exert effort is important; we can shorten the notations and denote $t_m = t(1^m, 0^{n-m})$, $\Delta_m = t_m - t_{m-1}$, $p_m = \frac{c}{\Delta_m}$ and $u_m = t_m \cdot (v - m \cdot p_m)$.

3. Warmup: Analysis of anonymous AND and OR

As a warmup we discuss two families of simple anonymous technologies, the family of AND technologies and the family of OR technologies. In these technologies each agent has his own "task," and each agent succeeds or fails in his task independently. The project's success depends on the set of successful sub-tasks: in an AND technology all agents need to succeed for the project to succeed, whereas in an OR technology it is enough that at least one agent succeeds for the project to succeed (it fails only if all agents fail). These two are extreme technologies: an AND technology exhibits only complementarities between the agents, whereas in an OR technology agents are perfect substitutes. We next present these technologies formally.

Fix two parameters γ and δ such that $0 < \gamma < \delta < 1$. An agent succeeds in his task with probability γ when he does not exert effort, and he succeeds with probability δ if he exerts effort.

In an anonymous AND technology with n agents, if m agents exert effort then $t_m = \delta^m \cdot \gamma^{n-m}$. E.g., for two agents, the AND success function is given by $t_0 = \gamma^2$, $t_1 = \delta \cdot \gamma$, and $t_2 = \delta^2$.

In an anonymous OR technology with n agents, if m agents exert effort then $t_m = 1 - (1 - \delta)^m (1 - \gamma)^{n-m}$. E.g. for two agents, the OR success function is given by $t_0 = 1 - (1 - \gamma)^2$, $t_1 = 1 - (1 - \delta)(1 - \gamma)$, and $t_2 = 1 - (1 - \delta)^2$.

To give the reader some intuition we next present a direct and complete analysis of AND and OR technologies for two agents, in the case where $c = 1$ and the parameters are $\gamma = 1 - \delta = 1/4$.

Example 3.1 *(AND technology with two agents, $c = 1, \gamma = 1 - \delta = 1/4$).* We have $t_0 = \gamma^2 = 1/16$, $t_1 = \gamma(1 - \gamma) = 3/16$, and $t_2 = (1 - \gamma)^2 = 9/16$; thus $\Delta_1 = 1/8$ and $\Delta_2 = 3/8$. The principal has 3 possibilities: contracting with 0, 1, or 2 agents. Let us write down the expressions for his utility in these 3 cases:

- **0 Agents:** No agent is paid and thus the principal's utility is $u_0 = t_0 \cdot v = v/16$.
- **1 Agent:** This agent is paid $p_1 = c/\Delta_1 = 8$ on success and the principal's utility is $u_1 = t_1(v - p_1) = 3v/16 - 3/2$.
- **2 Agents:** Each agent is paid $p_2 = c/\Delta_2 = 8/3$ on success, and the principal's utility is $u_2 = t_2(v - 2p_2) = 9v/16 - 3$.

Notice that the option of contracting with one agent is always inferior to contracting either with both or with none, and will never be taken by the principal. The principal will contract with no agent when $v < 6$, with both agents when $v > 6$, and with either none or both when $v = 6$.

This should be contrasted with the observable-actions case in which the principal observes the individual actions. In this case, the principal can pay each agent exactly his cost in order to induce him to exert effort. In the observable-actions case, therefore, the principal simply optimizes the social welfare. In the example above, the principal will contract with both agents when $v \geqslant 4$, and none when $v \leqslant 4$. Thus for example, for $v = 6$ the utility maximizing decision (in the observable-actions case) would give a principal's utility of $6 \cdot 9/16 - 2 = 11/8$ while the principal's decision (in the hidden-actions case) would give him utility of $3/8$, giving a ratio of $11/3$.

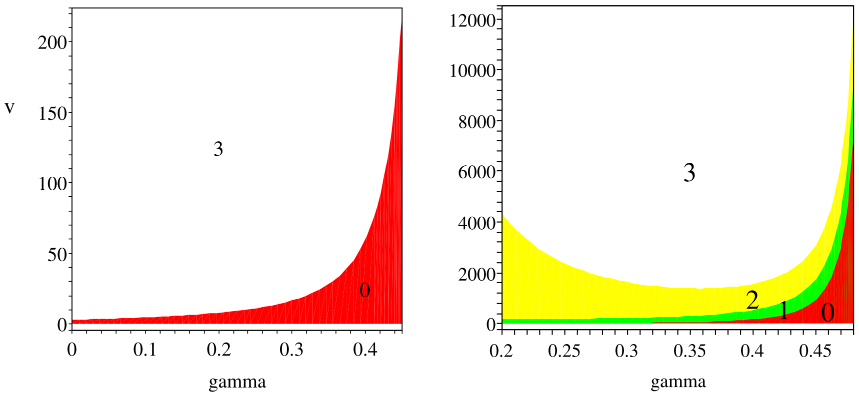

Fig. 1. Number of agents in the optimal contract of the AND (left) and OR (right) technologies with 3 agents, as a function of γ (horizontal axis) and v (vertical axis), where $\delta = 1 - \gamma$. AND technology: either 0 or 3 agents are contracted, and the transition value is monotonic in γ. OR technology: for any γ we can see all transitions.

It turns out that this ratio of $11/3$ at $v = 6$ is the social price of unobservability in this example, and it is obtained exactly at the transition point of the hidden-actions case. This point also yields the principal's price of unobservability, which is equal to the social POU in this example.

Example 3.2 *(OR technology with two agents, $c = 1$, $\gamma = 1 - \delta = 1/4$).* We have $t_0 = 1 - (1 - \gamma)^2 = 7/16$, $t_1 = 1 - \gamma(1 - \gamma) = 13/16$, and $t_2 = 1 - \gamma^2 = 15/16$ thus $\Delta_1 = 3/8$ and $\Delta_2 = 1/8$. Let us write down the expressions for the principal's utility in these three cases:

- **0 Agents:** No agent is paid and the principal's utility is $u_0 = t_0 \cdot v = 7v/16$.
- **1 Agent:** This agent is paid $p_1 = c/\Delta_1 = 8/3$ on success and the principal's utility is $u_1 = t_1(v - p_1) = 13v/16 - 13/6$.
- **2 Agents:** each agent is paid $p_2 = c/\Delta_2 = 8$ on success, and the principal's utility is $u_2 = t_2(v - 2p_2) = 15v/16 - 15$.

Now, contracting with one agent is better than contracting with none whenever $v > 52/9$ (and is equivalent for $v = 52/9$), and contracting with both agents is better than contracting with one agent whenever $v > 308/3$ (and is equivalent for $v = 308/3$); thus the principal will contract with no agent for $0 \leqslant v \leqslant 52/9$, with one agent for $52/9 \leqslant v \leqslant 308/3$, and with both agents for $v \geqslant 308/3$.

In contrast, in the observable-actions case the principal will make a single agent exert effort for $v > 8/3$, and the second one exert effort as well when $v > 8$.

It turns out that the social price of unobservability here is $19/13$, and is achieved at $v = 52/9$, which is exactly the transition point from 0 to 1 contracted agent in the hidden-actions case. The same is true for the principal's POU. It is not a coincidence that for both the AND and OR technologies the social POU and the principal's POU are obtained for v that is a transition point; this is a special case of the general result of Lemma 2.9.

We can already see a qualitative difference between the AND and OR technologies (even with 2 agents): in the first case either all agents are contracted or none is, whereas in the second case, for some intermediate range of values v, exactly one agent is contracted. Fig. 1 shows the same phenomenon for AND and OR technologies with 3 agents, and we next show that it holds for any number of agents.

Proposition 3.3. *For any anonymous AND technology,*[15] *in both the hidden- and observable-actions cases, there exists a value*[16] $v_* < \infty$ *such that for any* $v < v_*$ *it is optimal to contract with no agent, for* $v > v_*$ *it is optimal to contract with all n agents, and for* $v = v_*$ *both contracts (with none or all agents) are optimal.*

The property of a single transition in the AND technology occurs in both the hidden-actions and the observable-actions cases, where the transition occurs at a smaller value of v in the observable-actions case. In Section 4 we present a general characterization of technologies with a single transition in the hidden-actions and the observable-actions cases. The above proposition is a special case of the analysis given in Section 4.

Next we consider anonymous OR technologies and show that any such technology exhibits all n transitions.

Proposition 3.4. *For any anonymous OR technology,*[17] *in both the hidden- and observable-actions cases, there exist finite positive values* $v_1 < v_2 < \cdots < v_n$ *such that for any* v *s.t.* $v_k < v < v_{k+1}$*, contracting with exactly k agents is optimal (for* $v < v_1$*, no agent is contracted, and for* $v > v_n$*, all n agents are contracted). For* $v = v_k$*, the principal is indifferent between contracting with* $k - 1$ *or k agents.*

This characterization is also a direct corollary of a more general characterization given in Section 4.

4. What determines the transitions?

In this section we analyze properties of general and anonymous technologies that determine the transition structure of these technologies, in both the hidden-actions case and the observable-actions case. In particular, we are interested in the extreme cases, i.e., the case of only one transition point, where the optimal contract changes from the empty set to all agents (and any other contract is never optimal), and the case of a maximal number of distinct transition points (which implies a maximal number of optimal contracts that are not equivalent[18]). For anonymous technologies the maximal number of transitions is n. For general technologies the maximal number of transitions is n in the observable-actions case, but can be larger in the hidden-actions case. For all cases but the last we present exact characterizations of the properties of the technologies that will ensure the relevant transition structure.

For general technologies in the hidden-actions case, the number of transition points can potentially be exponential in the number of agents n (potentially, $2^n - 1$). It is easy to observe that

[15] Anonymous AND technology with any number of agents n and any parameters γ and δ such that $1 > \delta > \gamma > 0$, and any identical cost c.

[16] v_* is a function of n, γ, δ, and c.

[17] Anonymous OR technology with any number of agents n and any parameters γ and δ such that $1 > \delta > \gamma > 0$, and any identical cost c.

[18] For example, in an anonymous technology all contracts with the same number of agents are essentially the same and are counted only once.

it cannot be the case that each of the 2^n possible contracts is the unique optimal contract for some value (which implies $2^n - 1$ distinct transition points). It suffices to observe that out of all contracts with a single agent, only a contract with an agent i for which $t(\{i\}) \geqslant t(\{j\})$ for any agent j can be optimal. Given this observation a natural question is what is the maximal number of transition points that a technology can have in the hidden-actions case. As far as we know this question is still open, but it was shown by [5] that the number can be very large, exponential in n, and [3] showed that even a non-anonymous OR technology can have an exponential number of transitions.[19]

We consider general technologies in Section 4.1 and anonymous technologies in Section 4.2. In Appendix B.3 we discuss the relationship between the different properties, in particular we show that even for anonymous technologies it is not necessarily the case that one transition in the observable-actions case implies one transition in the hidden-actions case, or vice versa. Additionally, for anonymous technologies it is not necessarily the case that n transitions in the observable-actions case implies n transitions in the hidden-actions case, or vice versa.

We refer the reader to Appendix B for all proofs omitted from this section.

4.1. General technologies

In this section we consider general technologies and study the properties that will ensure an extreme number of transitions in the hidden-actions and observable-actions cases.

We first present properties that ensure that there is only one transition point, and the optimal contract changes from the empty set to all agents (any other contract is never optimal).

Definition 4.1. A technology success function t exhibits (strict) **under-proportional contribution (UPC)** if for every non-empty set $S \subsetneq N$ it holds that

$$\frac{|S|}{|N|} > \frac{t(S) - t(\emptyset)}{t(N) - t(\emptyset)}.$$

Note that the UPC condition is equivalent to requiring that for any non-empty subset of agents $S \subsetneq N$ it holds that $\frac{t(N)-t(\emptyset)}{|N|} > \frac{t(S)-t(\emptyset)}{|S|}$. That is, UPC means that for any non-empty S, the average contribution of each of the $|S|$ agents is smaller than the average contribution of each agent when all agents ($|N| = n$ agents) exert effort. The UPC property is weaker than IRS, as any technology that exhibits IRS also exhibits UPC (Lemma B.9), but not vice versa (Proposition B.10).

For a technology (t, c) let $Q(S)$ be the total expected payment in the best contract that induces an equilibrium in which the agents in S exert effort (i.e., the minimal cost of implementing a contract with the set of agents S), in the hidden-actions case. As we have seen, $Q(S) = t(S) \sum_{i \in S} \frac{c}{\Delta_i(S_{-i})}$.

Definition 4.2. A technology (t, c) exhibits (strict) **over-payment (OP)** if for every non-empty subset $S \subsetneq N$, it holds that

$$\frac{Q(S)}{Q(N)} > \frac{t(S) - t(\emptyset)}{t(N) - t(\emptyset)}.$$

[19] In the proof (in Appendix B) we present a necessary and sufficient condition for a list of sets of agents to be the list of optimal contracts as the value varies; the challenge is to find the largest such list.

Intuitively, the over-payment condition requires that the proportional increase in success probability when moving from a set of agents S to the set of all agents N (which is $\frac{t(S)-t(\emptyset)}{t(N)-t(\emptyset)}$), be smaller than the proportional expected payment to the agents in S with respect to the expected payment to the set N (which is $\frac{Q(S)}{Q(N)}$). Alternatively, for any non-empty set $S \neq N$, the ratio of its expected payment to its increase in success probability (given by $\frac{Q(S)}{t(S)-t(\emptyset)}$) is greater than that ratio with respect to the set N. Theorem 4.3 below establishes that if for any set S, the agents in S need to be paid "overproportionally," the principal will contract either with 0 agents or with all n agents.

The following theorem provides a characterization of technologies that exhibit a single transition (only the empty contract or the contract with all agents can be optimal).

Theorem 4.3. *A technology (t, c) has a single transition in the hidden-actions case if and only if it exhibits over-payment (OP). In addition, a technology (t, c) has a single transition in the observable-actions case if and only if it exhibits under-proportional contribution (UPC).*

The proofs of both results are very similar (relying on the fact that the OP condition reduces to UPC if the hidden-actions payment function $Q(s) = t(S)\sum_{i\in S}\frac{c}{\Delta_i(S_{-i})}$ is replaced by the observable-actions payment function $Q(S) = c|S|$), such that in each case the stated condition ensures that in the transition point between the empty contract and the contract with all agents, the utility of the principal from any other contract is smaller. That is sufficient to ensure that any other contract is never optimal.

Single transition is closely related to complementary. IRS implies UPC and thus implies a single transition in the observable-actions case. Although IRS is not sufficient to imply a single transition in the hidden-actions case (Proposition B.10) a strengthening of the IRS property (defined below) does imply OP and thus a strong version of complementarity implies a single transition in the hidden-actions case. Lemma B.12 states that any technology (t, c) that exhibits *strong IRS* exhibits *OP*. We next define *strong IRS*.

Definition 4.4. A technology success function t exhibits **strong increasing returns to scale (strong IRS)** if for every i and every $a_{-i} < b_{-i}$ it holds that $\frac{\Delta_i(a_{-i})}{t(1,a_{-i})} \leqslant \frac{\Delta_i(b_{-i})}{t(1,b_{-i})}$.

Note that this is a stronger condition than *IRS*. *IRS* only requires that $\Delta_i(a_{-i}) < \Delta_i(b_{-i})$, whereas the new condition requires that $\Delta_i(a_{-i}) \cdot t(1, b_{-i}) \leqslant \Delta_i(b_{-i}) \cdot t(1, a_{-i})$, and as $a_{-i} < b_{-i}$ it holds that $t(1, a_{-i}) < t(1, b_{-i})$; thus strong IRS is harder to satisfy.

One can also ask what is the relation between the over-payment property (OP) and the under-proportional contribution (UPC) property. In Proposition B.10 we prove that even for anonymous technologies, neither implies the other. On the other hand, it is possible for a technology to satisfy both; for example, this is the case for any anonymous AND technology (Observation 4.11), and thus any such technology has a single transition in both the observable- and hidden-actions cases.

We next present a property that ensures that there are n distinct transition points in the case of observable actions. Note that in the observable-actions case (under the identical costs assumption) the cost of any contract with k agents is the same; thus of all contracts with exactly k agents only contracts that maximize the success probability can be optimal.

For any $k \in \{0, 1, \dots, n\}$, let S_k^* be a set of size k such that $t(S_k^*) \geqslant t(S_k)$ for any set S_k of size k.

Definition 4.5. A technology success function t exhibits (strict) **decreasing returns to maximal scale (DRMS)** if for every $k \in \{1, 2, \ldots, n-1\}$ it holds that

$$t(S_k^*) - t(S_{k-1}^*) > t(S_{k+1}^*) - t(S_k^*).$$

Theorem 4.6. *A technology* (t, c) *has all n distinct transitions in the observable-actions case if and only if it exhibits decreasing returns to maximal scale* (*DRMS*).

Note that if for some k there are multiple contracts of size k with the same maximal success probability $t(S_k^*)$, then for any value for which one such contract is optimal, any such contract is optimal. If we add the assumption that for every k there is a unique set with maximal success probability $t(S_k^*)$, then n distinct transitions means that there are only $n+1$ contracts that can be optimal, and for each $k \in \{0, 1, \ldots, n\}$ the contract S_k^* is the unique optimal contract for any value between the transition point from S_{k-1}^* to S_k^* and the transition point from S_k^* to S_{k+1}^*.

4.2. Anonymous technologies

In this section we consider anonymous technologies and study properties that ensure an extreme number of transitions in the hidden-actions and observable-actions cases. Clearly, anonymous technologies have at least one transition point and at most n such points. For the single-transition point the properties are essentially the specification of the general properties (UPC and OP) to the anonymous case. For n transition points, in the observable-actions case the property is a specification of DRMS to the anonymous case, and it turns out to be the standard decreasing returns to scale (DRS) for the anonymous case. The main additional result in this section is a characterization of n transitions for anonymous technologies in the hidden-actions case.

Recall that for anonymous technology t we denote by t_k the success probability of the project if k agents exert effort. The under-proportional contribution (UPC) property of Definition 4.1 can now be written for an anonymous technology t as follows.

An anonymous technology success function t exhibits (strict) *under-proportional contribution* (*UPC*) if for every $k \in \{1, 2, \ldots, n-1\}$ it holds that

$$\frac{k}{n} > \frac{t_k - t_0}{t_n - t_0}.$$

When we talk about a property that holds for anonymous technologies we sometimes make it explicit in the notation in the following way. The under-proportional contribution (*UPC*) for anonymous technology will be denoted by UPC_{anon}. Similar notations will be used for any other properties on anonymous technologies.

For an anonymous technology (t, c) let $Q_k = \frac{c \cdot k \cdot t_k}{t_k - t_{k-1}}$ be the total expected payment in the best contract for which there exists an equilibrium with k agents exerting effort (the minimal cost of implementing a contract with k agents). The over-payment (OP) property of Definition 4.2 can now be written for an anonymous technology t as follows.

An anonymous technology exhibits (strict) *over-payment* (OP_{anon}) if for every $k \in \{1, 2, \ldots, n-1\}$, it holds that

$$\frac{Q_k}{Q_n} > \frac{t_k - t_0}{t_n - t_0}.$$

The following corollary is immediate from Theorem 4.3 and the above definitions.

Corollary 4.7. *An anonymous technology (t, c) has a single transition in the hidden-actions case if and only if it exhibits over-payment (OP_{anon}). In addition, an anonymous technology (t, c) has a single transition in the observable-actions case if and only if it exhibits under-proportional contribution (UPC_{anon}).*

Next we consider the properties that ensure n distinct transition points. We first consider the observable-actions case. Clearly, for an anonymous technology every set of size k has the same success probability. Thus the decreasing returns to maximal scale (DRMS) when considered in the anonymous case is equivalent to the requirement that for every $k \in \{1, 2, \ldots, n-1\}$ it holds that

$$t(t_k) - t(t_{k-1}) > t(t_{k+1}) - t(t_k)$$

which is exactly the standard decreasing returns to scale (DRS) for an anonymous technology. Thus the following corollary is immediate from Theorem 4.6 and the above observation.

Corollary 4.8. *An anonymous technology (t, c) has all n distinct transitions in the observable-actions case if and only if it exhibits decreasing returns to scale (DRS).*

We next discuss a property that ensures n distinct transition points in the hidden-actions case.

Definition 4.9. An anonymous technology (t, c) exhibits (strictly) **increasing relative marginal payment (**$IRMP_{anon}$**)** if for every $k \in \{1, 2, \ldots, n-1\}$ it holds that

$$\frac{Q_{k+1} - Q_k}{t_{k+1} - t_k} > \frac{Q_k - Q_{k-1}}{t_k - t_{k-1}}.$$

This condition considers the proportional increase in payment with respect to the increase in success probability ($\frac{Q_k - Q_{k-1}}{t_k - t_{k-1}}$), when the number of contracting agents is increased by one, from $k-1$ to k. The $IRMP_{anon}$ condition requires that any additional agent have a larger effect than its predecessor on that ratio. The $IRMP_{anon}$ condition ensures n distinct transition points in the hidden-actions case.

Theorem 4.10. *An anonymous technology (t, c) has all n transitions in the hidden-actions case if and only if it exhibits increasing relative marginal payment ($IRMP_{anon}$).*

One can ask what is the relation between the decreasing returns to scale (DRS_{anon}) and increasing relative marginal payment ($IRMP_{anon}$) properties for anonymous technologies. In Proposition B.14 we prove that neither implies the other. On the other hand, it is possible for a technology to satisfy both; for example, we show that this is the case for any anonymous OR technology (Observation 4.12); thus any such technology has n transitions in both the observable- and hidden-actions cases.

Now that we have a complete characterization of the properties that ensure extreme transitions, we can go back and relate these properties to the two technologies discussed in Section 3, the anonymous AND and OR technologies. We state the following two results.

Observation 4.11. *Any anonymous AND technology exhibits both under-proportional contribution (UPC_{anon}) and over-payment (OP_{anon}), and thus has a single transition in both the hidden-actions and observable-actions cases.*

Observation 4.12. *Any anonymous OR technology exhibits both decreasing returns to scale (DRS_{anon}) and increasing relative marginal payment ($IRMP_{anon}$), and thus has all n transitions in both the hidden-actions and observable-actions cases.*

5. Agent selection

Given two agents, which one of them should be contracted? More generally, what determines the selection of the agents to be incentivized in an optimal contract? We define two agent properties, both of which induce a relation on the agents and reflect the agents' indispensability to the production process, which in turn determines the agent's selection. The proofs of this section are deferred to Appendix C.

One agent is said to be more productive than another agent if by replacing the first by the second the project's success probability decreases.

Definition 5.1. Agent j is *more productive* than agent i if for all subsets $S \subset N \setminus \{i, j\}$ it holds that $t(S \cup \{j\}) > t(S \cup \{i\})$.

As mentioned above, productivity induces a relation on the agents. It is possible that for two agents, neither of them is more productive than the other.

Note that in the observable-actions case the principal always prefers contracting with a more productive agent to contracting with a less productive one.

Observation 5.2. *Fix a technology (t, c). In the observable-actions case if agent j is more productive than agent i then it holds that for every v, for every optimal contract S^*, if $i \in S^*$ then $j \in S^*$.*

This means that in the observable-actions case if j is more productive than i it is never the case that the principal decides to contract with i and not with j.

We next show that, unlike the observable-actions case, under hidden actions a less productive agent may be contracted in an optimal contract, while a more productive one will be left out.

Example 5.3. Consider the following technology with 3 agents. Let $\epsilon = 1/100$. $t(\emptyset) = 0$; $t(\{1\}) = 1 - 5\epsilon$; $t(\{2\}) = \epsilon$; $t(\{3\}) = \epsilon$; $t(\{1, 2\}) = 1 - 5\epsilon + \epsilon^2$; $t(\{1, 3\}) = 1 - \epsilon + \epsilon^2$; $t(\{2, 3\}) = 1 - \epsilon$; $t(\{1, 2, 3\}) = 1 - \epsilon + \epsilon^2 + \epsilon^3$. Assume that the cost of effort is $c = 1$ for every agent.

Note that agent 1 is more productive than agent 2 as $t(\{1\}) > t(\{2\})$ and $t(\{1, 3\}) > t(\{2, 3\})$. Nevertheless, we show that in the hidden-actions case it is possible that the principal contracts with $\{2, 3\}$, leaving 1, who is more productive than 2, out of the contract. This is in contrast to the observable-actions case (Observation 5.2).

Assume that $v = 200$. It is easy to see that a contract with agent 1 is better than any contract with any other single agent and it gives the principal's utility of $u(\{1\}, 200) = 189$. Contracting with $\{2, 3\}$ gives the principal a higher utility of $u(\{2, 3\}, 200) = 195\frac{48}{49}$. The contract $\{1, 3\}$ results in lower utility as agent 3 should be paid $\frac{1}{1-\epsilon+\epsilon^2-(1-5\epsilon)} = \frac{1}{4\epsilon+\epsilon^2} > 24$; thus the utility of the principal is at most $200 - 24 = 176 < 195\frac{48}{49}$. Similarly, agent 2 must receive a huge payment of at least $1/\epsilon^2 = 10\,000$ in both the contracts $\{1, 2\}$ and $\{1, 2, 3\}$, which yields negative utility. We conclude that the contract $\{2, 3\}$ is optimal for $v = 200$.

Taking a closer look at the example, we find that while agent 1 is more productive than agent 2, for some value the contract $\{2, 3\}$ yields higher utility than the contract $\{1, 3\}$. The explanation behind this phenomenon is that the marginal contribution of agent 3 in the contract $\{1, 3\}$ is very small, which results in a very high payment to agent 3 in that contract. In contrast, agent 3's marginal contribution in the contract $\{2, 3\}$ is much higher, and thus the contract $\{2, 3\}$ results in higher utility to the principal although agent 2 is less productive than agent 1. We conclude that productivity alone does not guarantee precedence. We next introduce a condition that combined with productivity ensures precedence of one agent over another.

Definition 5.4. Fix two agents i and j and a subset $S \subset N \setminus \{i, j\}$. Define $S_i = S \cup \{i\}$ and $S_j = S \cup \{j\}$. Agent j is a *better productivity initiator* than agent i if for all subsets $S \subset N \setminus \{i, j\}$ and for every agent $k \notin S$ it holds that

$$t\big(S_j \cup \{k\}\big) - t(S_j) > t\big(S_i \cup \{k\}\big) - t(S_i).$$

This means that agent k has a larger influence on the success of the project when i is replaced by j. Clearly it is possible that j is not a better productivity initiator than i, nor agent i a better productivity initiator than j. We are now ready to present a sufficient condition for a more productive agent to get precedence over a less productive one.

Theorem 5.5. *Fix a technology (t, c). In the hidden-actions case if agent j is more productive than agent i and also a better productivity initiator than i then it holds that for every v and for every optimal contract S^*, if $i \in S^*$ then $j \in S^*$.*

Example 5.3 above shows that there exists a technology (t, c) in which agent j (agent 1) is more productive than i (agent 2) while i is a better productivity initiator than j, and for that technology there is an optimal contract (for some value) that includes i but not j. The same example also shows that although i is a better productivity initiator than j it is possible that for some value an optimal contract contains j but not i (as i is less productive). In the example this is the set that only includes j (agent 1) which is optimal for value of 20 (the principal utility is 18 for the set $\{1\}$ and lower for any other set). We conclude that by themselves, neither higher productivity nor better productivity initiation ensures precedence of one agent over the other, but the combination of the two does ensure such precedence. This is in contrast to the observable-actions case, in which the productivity relation alone determines precedence of one agent over another (see Observation 5.2).

To demonstrate the usefulness of Theorem 5.5 let us consider non-anonymous AND technologies with agent i having success probability of δ_i in his task when exerting effort, and $\gamma_i < \delta_i$ when he does not. The probability of project success is the product of the success probabilities of the tasks. Sort the agents in decreasing order of δ_i/γ_i and assume that there are no ties. One can easily verify that this order also corresponds to the productivity and productivity initiation order over the agents. We conclude that there are at most $n + 1$ different optimal contracts for any such AND technology, and that we can compute the optimal contract for a value by simply checking all possible $n + 1$ prefixes of the agents according to the above order.

6. Price of unobservability (POU)

In this section we consider the social POU and the principal's POU, which quantifies the worst-case loss due to the lack of observability of the agents' actions. For the formal definitions of the social POU and the principal's POU we refer the reader to Section 2.3. We find that this loss is strongly affected by whether agents substitute or complement each other. The proofs of this section are deferred to Appendix D.

We first show that the social POU and the principal's POU of any technology that exhibits Decreasing Returns to Scale (DRS) are always upper bounded by a linear function of the number of agents.

Theorem 6.1. *For any technology (t, c) with n agents that exhibits DRS, it holds that $POU_S \leqslant POU_P \leqslant 2n$.*

We also find that the upper bound for the POU_P given in the last theorem is (almost) tight. In fact, this bound is obtained already with an anonymous DRS technology.

Theorem 6.2. *There exists an anonymous technology with n agents that exhibits DRS for which $POU_P \geqslant n/2$.*

While the price of unobservability of technologies that exhibit DRS is always bounded from above by a linear function of the number of agents, this is not the case for IRS technologies. In particular, for the family of IRS technologies, the price of unobservability is not at all bounded, even for anonymous technologies. This is cast in the following theorem.

Theorem 6.3. *The social POU and the principal's POU are not bounded from above across the family of IRS technologies (even anonymous ones).*

Proof. The following lemma (whose proof is deferred to Appendix D) fully characterizes the social POU and the principal's POU of anonymous technology that exhibits both UPC and OP.

Lemma 6.4. *For any anonymous technology (t, c) that exhibits both UPC and OP it holds that*

$$POU_S = POU_P = 1 + \frac{t_{n-1}}{t_0} - \frac{t_{n-1}}{t_n}$$

and both are obtained at a transition point of the hidden-actions case.

To show that the social POU and the principal's POU are not bounded from above across the family of IRS technologies, it is sufficient to prove that they are not bounded across the family of anonymous AND technologies, as any such technology exhibits IRS (Observation B.15).

Based on Observation 4.11, any anonymous AND technology exhibits both UPC and OP. For any anonymous AND technology, with cost c and parameters $\delta > \gamma > 0$, it holds that $\frac{t_{n-1}}{t_0} = \frac{\delta^{n-1}\cdot\gamma}{\gamma^n} = (\frac{\delta}{\gamma})^{n-1}$, and $\frac{t_{n-1}}{t_n} = \frac{\delta^{n-1}\cdot\gamma}{\delta^n} = \frac{\gamma}{\delta}$. Substituting the last values in the expression given in Lemma 6.4, we get that for such an anonymous AND technology,

$$POU_P = POU_S = \left(\frac{\delta}{\gamma}\right)^{n-1} + 1 - \frac{\gamma}{\delta}.$$

It follows that the social POU and the principal's POU approach ∞ either if $\gamma \to 0$ (for any given $n \geqslant 2$ and δ bounded from 0, in particular when $\delta = 1 - \gamma > 1/2$) or if $n \to \infty$ (for any fixed $\delta > \gamma > 0$). We conclude that the social POU and the principal's POU are not bounded across the family of anonymous AND technologies (for various n, γ). Since any anonymous AND technology exhibits IRS, it follows that the social POU and the principal's POU are not bounded across the family of IRS technologies. □

We observe a qualitative difference between technologies that exhibit complementarities and substitutability. While in the family of IRS technologies both POU_P and POU_S are not bounded from above, in the family of DRS technologies both are always bounded from above by a linear function of the number of the agents. These results have direct implications on the effectiveness of investing in monitoring to obtain more information about the agents' effort. On the one hand, monitoring provides more information and may improve the principal's utility. On the other hand, monitoring is often very costly. Therefore, a principal should conduct a cost–benefit analysis to determine whether investing in monitoring is worthwhile. If there is a large loss due to hidden actions (as may be the case in technologies that exhibit complementarities), the principal may be better off investing in technologies to better monitor the agents. If, however, the loss due to hidden actions is relatively small (as in technologies that exhibit substitutabilities), the principal may be better off conditioning the payment on the final outcome alone without investing in monitoring. Our results can help a principal quantifying this tradeoff to reach the right decision.

7. Conclusions

The optimal design of incentives in teams with moral hazard heavily depends on the structure of the team technology, with substitution and complementarity playing crucial roles. In this paper we paid attention to two aspects of the design of teams:

1. The size and structure of the selected group and its transition as the value of the project changes.
2. The principal's loss of profit due to the moral hazard.

We have found that the properties of complementarity and substitutability of the production technology induce polar results for both the design issues discussed above. Our analysis and results concerning the difference between complementarity and substitution in the design of optimal teams have some testable implications: under complementarity slight changes in the principal's benefits from the team's success may result in massive changes in the team's size. Under substitution, in contrast, as the principal's benefit goes up (or down) the optimal group size varies only gradually. The principal's loss of profit due to moral hazard is greater under complementarity, which suggests that it is more likely to see the principal investing in monitoring for such technologies. We view many of the results in this paper, and in particular the analysis in Section 6, as a starting point to an approach that endogenizes the information structure faced by the principal in contrast to the standard approach where this structure is assumed exogenously fixed and uncontrolled.

Appendix A. Model and preliminaries

Lemma 2.5 *(Monotonicity lemma). For any technology* $(t, \vec{c})$*, in both the hidden-actions and the observable-actions cases, all the following are monotonically non-decreasing with the value*: (i) *the expected utility of the principal at the optimal contracts,* (ii) *the success probability of the optimal contracts, and* (iii) *the expected payment of the optimal contracts.*

Proof. Suppose the sets of agents S_1 and S_2 are optimal in v_1 and $v_2 < v_1$, respectively. Let $Q(S)$ denote the expected total payment to all agents in S in the case that the principal contracts with the set S and the project succeeds (for the hidden-actions case, $Q(S) = t(S) \cdot \sum_{i \in S} \frac{c_i}{t(S)-t(S \setminus i)}$, while for the observable-actions case $Q(S) = \sum_{i \in S} c_i$). The principal's utility is a linear function of the value, $u(S, v) = t(S) \cdot v - Q(S)$. As S_1 is optimal at v_1, $u(S_1, v_1) \geqslant u(S_2, v_1)$, and as $t(S_2) \geqslant 0$ and $v_1 > v_2$, $u(S_2, v_1) \geqslant u(S_2, v_2)$. We conclude that $u(S_1, v_1) \geqslant u(S_2, v_2)$; thus the utility is monotonic non-decreasing in the value.

Next we show that the success probability is monotonic non-decreasing in the value. S_1 is optimal at v_1; thus

$$t(S_1) \cdot v_1 - Q(S_1) \geqslant t(S_2) \cdot v_1 - Q(S_2).$$

S_2 is optimal at v_2; thus

$$t(S_2) \cdot v_2 - Q(S_2) \geqslant t(S_1) \cdot v_2 - Q(S_1).$$

Summing these two equations, we get that $(t(S_1) - t(S_2)) \cdot (v_1 - v_2) \geqslant 0$, which implies that if $v_1 > v_2$ then $t(S_1) \geqslant t(S_2)$.

Finally, we show that the expected payment is monotonic non-decreasing in the value. As S_2 is optimal at v_2 and $t(S_1) \geqslant t(S_2)$, we observe that:

$$t(S_2) \cdot v_2 - Q(S_2) \geqslant t(S_1) \cdot v_2 - Q(S_1) \geqslant t(S_2) \cdot v_2 - Q(S_1),$$

or equivalently, $Q(S_2) \leqslant Q(S_1)$, which is what we wanted to show. □

Observation 2.8. *For any technology t it holds that $POU_P(t) \geqslant POU_S(t)$.*

Proof. The numerators of the two expressions are identical. Therefore, it is sufficient to show that for any v, it holds that

$$t\big(S^*(v)\big) \cdot v - \sum_{i \in S^*(v)} c_i \geqslant t\big(S^*(v)\big) \cdot \left(v - \sum_{i \in S^*(v)} \frac{c_i}{t(S^*(v)) - t(S^*(v) \setminus \{i\})}\right),$$

but this is true since the principal's utility in the hidden-actions case is at most the social welfare (as each agent has non-negative utility). □

In what follows, we show that the price of unobservability is always obtained at a transition point.

Lemma 2.9. *For any given technology $(t, \vec{c})$ the social price of unobservability is obtained at some value v which is a transition point, of either the hidden-actions or the observable-actions case. The same holds for the principal's price of unobservability.*

Proof. For technology $(t, \vec{c})$ let $f(v)$ be the social welfare ratio at v, that is, $f(v) = \frac{t(S^*_{oa}) \cdot v - \sum_{i \in S^*_{oa}} c_i}{t(S^*) \cdot v - \sum_{i \in S^*} c_i}$. By definition the social POU is $POU_S(t, \vec{c}) = Sup_{v>0} f(v)$. By the monotonicity lemma (Lemma 2.5), in both the hidden-actions and observable-actions cases, contracting with no agent can be optimal up to some value, and never optimal for larger values. Additionally, once contracting with all agents is optimal, no other contract can be optimal for any larger value. Let $\underline{v}$ be the lowest value transition point, from contracting with no agent to contracting with some agents, of either the hidden-actions or observable-actions case. Let $\bar{v}$ be the highest value transition point, from contracting with some agents to contracting with all agents, of either the hidden-actions or observable-actions case. Note that for value $v \leqslant \underline{v}$ no agent is contracted in both the hidden-actions and observable-actions cases, and for value $v \geqslant \bar{v}$ all agents are contracted in both cases. Thus for $v \leqslant \underline{v}$ and for $v \geqslant \bar{v}$, $f(v) = 1$. We conclude that if the social POU is obtained, it happens for a finite positive value $v \in [\underline{v}, \bar{v}]$.

We can assume w.l.o.g. that ties between optimal sets are broken in a consistent way (as we only care about the welfare of the principal). By Lemma 2.5, we can partition the $[\underline{v}, \bar{v}]$ interval to at most 2^n segments; in each the optimal contract in the hidden-actions case is fixed. Similarly, each of these segments can be partitioned to at most 2^n segments; in each the optimal contract for the observable-actions case is fixed. We conclude that there exists a finite partition of the $[\underline{v}, \bar{v}]$ interval to segments, such that the end points of the segments are transition points, and on each segment there is one contract that is optimal for the entire segment for both the hidden-actions and observable-actions cases.

To complete the proof, we use the following lemma (see proof below) which shows that for each segment, the supremum of $f(v)$ is obtained at an end point of the segment.

Lemma A.1. *Let $f(x) = \frac{a \cdot x - b}{c \cdot x - d}$ be a function of x, and assume that $c > 0$. Let $\bar{x} \geqslant \underline{x} > 0$ be two points for which $c\underline{x} - d > 0$. Then the supremum of f on the range $[\underline{x}, \bar{x}]$ is obtained at either $\underline{x}$ or $\bar{x}$. Additionally, if $a = c > 0$, $d > b$ and for some $\bar{x} > 0$ it holds that $a\bar{x} - d > 0$ then the supremum of f on $[\bar{x}, \infty)$ is obtained at $\bar{x}$.*

Proof. As f is a continuous function (recall that $cx - d > 0$ on the range as $c > 0$ and $c\underline{x} - d > 0$) on a compact range, its supremum is obtained.

In order to find the maximum of f, we take the first derivative and equate to zero:

$$\frac{\partial f}{\partial x} = \frac{a(cx-d) - c(ax-b)}{(cx-d)^2} = \frac{bc - ad}{(cx-d)^2} = 0$$

which holds if and only if $bc = ad$. As this equality is independent of x, it either holds for any x (and in particular for $\bar{x}$ and $\underline{x}$), or for no x. If it holds for no x, then the maximum must be obtained at either $\underline{x}$ or $\bar{x}$.

If $a = c > 0$, $d > b$, and for $\bar{x}$, $a\bar{x} - d > 0$, then f is continuous on $[\bar{x}, \infty)$. Additionally, $\frac{\partial f}{\partial x} = \frac{bc-ad}{(cx-d)^2} = \frac{a \cdot (b-d)}{(ax-d)^2} < 0$ for any $x \geqslant \bar{x}$. Thus the function f monotonically decreases on $[\bar{x}, \infty)$ and the supremum of f is obtained at $\bar{x}$. □

With this we are ready to complete the proof. On each of the segments mentioned above, $f(v)$ satisfies the conditions of the first part of Lemma A.1; thus its supremum is obtained at an end point of the segment. The global supremum (over all segments) is obtained as the maximum of finitely many maximal numbers is obtained, one in each segment. Therefore, the lemma holds for the social POU.

Let $g(v) = \frac{t(S^*_{oa}) \cdot v - \sum_{i \in S^*_{oa}} c_i}{t(S^*) \cdot v - \sum_{i \in S^*} \frac{c_i}{t(S^*) - t(S^* \setminus \{i\})}}$ be the principal's utility ratio at v. By definition the principal's POU is $POU_P(t, \vec{c}) = Sup_{v>0} g(v)$. For any $v \leqslant \underline{v}$, $g(v) = 1$. Additionally, $g(v)$ satisfies the conditions of the second part of Lemma A.1; thus we need not care about values $v \geqslant \bar{v}$. Finally, for values $[\underline{v}, \bar{v}]$, the same arguments that we use to prove that the POU_S is obtained at a transition point also hold for $g(v)$, and therefore the lemma holds for the principal's POU as well. □

Appendix B. What determines the transitions?

B.1. One transition

The following lemma presents necessary and sufficient conditions for a single transition. A single transition means that up to some value the empty set is the only optimal contract, and from that value contracting with all the agents is the only optimal contract (and at the value both contracts are optimal, and no other contract is optimal). The characterization holds for both the hidden-actions and the observable-actions cases, when in each case it should be applied with its corresponding payment function. The lemma implies Theorem 4.3 since the condition is exactly OP for the hidden-actions setting, and it is UPC in the observable-actions setting (as in the identical-costs case the ratio of the payments $\frac{Q(S)}{Q(N)}$ is exactly the ratio of the sets sizes $\frac{|S|}{|N|}$).

Lemma B.1. *Fix any technology* (t,c)*. In a given setting (observable actions or hidden actions) let* $Q(S)$ *be the total expected payment to all agents in the best contract that induces* S *to exert effort.*

There is a single transition if and only if for every non-empty subset $S \subsetneq N$*, it holds that*

$$\frac{Q(S)}{Q(N)} > \frac{t(S)-t(\emptyset)}{t(N)-t(\emptyset)}.$$

Proof. Let $u(S,v) = t(S)\cdot v - Q(S)$ be the utility of the principal with value v when optimally contracting with the set S (with $Q(S)$ being the total expected payment to all agents). Let $v_{0,n}$ be the value in which the principal is indifferent between the contract with no agent, and the contract with all n agents. It holds that $t(\emptyset)\cdot v_{0,n} = t(N)\cdot v_{0,n} - Q(N)$, or equivalently, $v_{0,n} = \frac{Q(N)}{t(N)-t(\emptyset)}$.

We first show that for every non-empty set $S \subsetneq N$

$$\frac{Q(S)}{Q(N)} > \frac{t(S)-t(\emptyset)}{t(N)-t(\emptyset)} \quad \Leftrightarrow \quad u(N,v_{0,n}) > u(S,v_{0,n}).$$

Indeed

$$\begin{aligned} u(N,v_{0,n}) > u(S,v_{0,n}) \quad &\Leftrightarrow \quad t(N)\cdot v_{0,n} - Q(N) > t(S)\cdot v_{0,n} - Q(S) \\ &\Leftrightarrow \quad \big(t(N)-t(S)\big)\cdot v_{0,n} > Q(N)-Q(S). \end{aligned}$$

As $v_{0,n} = \frac{Q(N)}{t(N)-t(\emptyset)}$ we can substitute and get

$$\begin{aligned} u(N,v_{0,n}) > u(S,v_{0,n}) \quad &\Leftrightarrow \quad \big(t(N)-t(S)\big)\cdot \frac{Q(N)}{t(N)-t(\emptyset)} > Q(N)-Q(S) \\ &\Leftrightarrow \quad \frac{Q(S)}{Q(N)} > \frac{t(S)-t(\emptyset)}{t(N)-t(\emptyset)} \end{aligned}$$

which is what we wanted to prove.

Given the above observation, to complete the proof of the lemma it is sufficient to show that a technology t has a single transition (from $\emptyset$ to N) if and only if $u(\emptyset,v_{0,n}) = u(N,v_{0,n}) > u(S,v_{0,n})$ for every non-empty set $S \subsetneq N$.

The "if" case: Assume that $u(N,v_{0,n}) > u(S,v_{0,n})$ for every non-empty set $S \subsetneq N$. For every set T it holds that $u(T,v)$ is linear in v, with a slope of $t(T)$. Since for every non-empty $S \subsetneq N$ it holds that $t(N) > t(S)$, the condition $u(N,v_{0,n}) > u(S,v_{0,n})$ implies that for every $v > v_{0,n}$ the contract N is the unique optimal contract. Additionally, since for every non-empty $S \subsetneq N$ it holds that $t(\emptyset) < t(S)$, the condition $u(\emptyset, v_{0,n}) > u(S, v_{0,n})$ implies that for every $v < v_{0,n}$ the contract $\emptyset$ is the unique optimal contract.

The "only if" case: Assume that t has a single transition from $\emptyset$ to N at $v_{0,n}$. As for any v any non-empty $S \subsetneq N$ is not optimal, S is not optimal at $v = v_{0,n}$ which means that $u(\emptyset, v_{0,n}) = u(N, v_{0,n}) > u(S, v_{0,n})$, as we needed to show. □

B.2. Multiple transitions

Our goal in this section is to provide necessary and sufficient conditions for n transition points, both in the general technologies observable-actions cases, and the anonymous technologies case (both for observable and hidden actions). The conditions will be derived from a more general result. We consider the problem of having some list of sets as the list of optimal contracts for a technology as the value varies. We provide a condition that ensures that this list will indeed be the list of optimal contracts. As in the case of a single transition, the characterization we provide holds for both the hidden-actions and the observable-actions cases, when in each case it should be applied with its corresponding payment function (and the derived utility function).

We begin by defining such a list of optimal contracts, which we call an *orbit*.

Definition B.2. Fix a technology (t, c) and assume that the utility of the principal with value v who is optimally contracting with the set S is $u(S, v)$.[20] A list of sets of agents $S_0 = \emptyset, S_1, S_2, \dots, S_m = N$ such that $t(S_k) < t(S_{k+1})$ for every $k \in \{0, 1, 2, \dots, m-1\}$ is called an *orbit* of size $m + 1$ of the technology (t, c) if the following hold:

- For any value $v \geqslant 0$ some set in the list is optimal (for any value v and any set S that is not in the list, there exists some k such that $u(S, v) \leqslant u(S_k, v)$).
- For every $k \in \{0, 1, 2, \dots, m\}$ there exists a value v_k such that S_k is optimal at v_k, and no set S_j for $j \neq k$ is optimal at v_k ($u(S_k, v_k) > u(S_j, v_k)$ for every $j \neq k$).

For a utility function $u(S, v)$ let $v_{i,j}$ be the value v in which the principal is indifferent between optimally contracting with the set S_i or with the set S_j.

The next observation shows that an orbit of size $m + 1$ implies that there are m transition points.

Observation B.3. *Fix a technology (t, c) and a utility function $u(S, v)$. If the list of sets of agents $S_0 = \emptyset, S_1, S_2, \dots, S_m = N$ is an orbit of size $m + 1$ of the technology (t, c), then the following hold*:

- *For every $k \in \{0, 1, 2, \dots, m\}$ the set S_k is optimal for every $v \in (v_{k-1,k}, v_{k,k+1})$, and for every $j \neq k$ the set S_j is not optimal at such a value v.*[21]
- *For every $k \in \{0, 1, 2, \dots, m\}$ both sets S_k and S_{k+1} are optimal at the value $v_{k,k+1}$.*

Proof. The second claim is by the definition of $v_{k,k+1}$ as the point of indifference between the sets S_k and S_{k+1}.

[20] The orbit depends on the utility function, which in turn depends on the payment function; thus the orbit is potentially different for the observable- and hidden-actions cases.

[21] With the convention that $v_{-1,0} = 0$ and $v_{m,m+1} = \infty$.

For the first claim, note that for each set T the utility $u(T, v)$ is linear in v, and the slope of the line is $t(T)$. Consider any value $v \in (v_{k-1,k}, v_{k,k+1})$ (note that it must be the case that $v_{k-1,k} < v_{k,k+1}$ as otherwise the orbit conditions are clearly violated). The fact that $t(S_{k+1}) > t(S_k) > t(S_{k-1})$, the definition of $v_{k-1,k}$ and $v_{k,k+1}$, as well as the linearity of the utility function $u(T, v)$ in v for every set T, together imply that $u(S_k, v) > u(S_{k+1}, v)$ and $u(S_k, v) > u(S_{k-1}, v)$.

Next we show that for $j \notin \{k-1, k, k+1\}$ it holds that $u(S_k, v) > u(S_j, v)$. We first consider the case where $j > k+1$. Assume on the contrary that for $j > k+1$ it holds that $u(S_k, v) \leqslant u(S_j, v)$. As both $u(S_{k+1}, v)$ and $u(S_j, v)$ are linear in v and $t(S_j) > t(S_{k+1})$ this implies that for any value $v' > v$ the set S_j is better than S_{k+1}. Moreover, it is clear that for any $v' < v_{k,k+1}$ (and thus for every $v' \leqslant v$) the set S_k is better than S_{k+1}. But this contradicts the second condition of the orbit as it means that for S_{k+1} there is no value v_{k+1} such that S_{k+1} is optimal at v_{k+1}, and no other set in the orbit is optimal at v_{k+1}. We conclude that $u(S_k, v) > u(S_j, v)$ for any $j > k+1$ and $v \in (v_{k-1,k}, v_{k,k+1})$. For the case of $j < k-1$, similar arguments show that if $u(S_k, v) \leqslant u(S_j, v)$ then there is no value v_{k-1} such that S_{k-1} is optimal at v_{k-1}, and no other set in the orbit is optimal at v_{k-1}. We conclude that $u(S_k, v) > u(S_j, v)$ for any $j \neq k$ and $v \in (v_{k-1,k}, v_{k,k+1})$.

Now, by the orbit's first condition, no set S that does not belong to the orbit is better than S_k for any such value v. Thus S_k is optimal for every $v \in (v_{k-1,k}, v_{k,k+1})$, and for every $j > k$ the set S_j is not optimal at such a value v. □

Thus, the above observation justifies the following definition.

Definition B.4. A technology (t, c) is said to have m *transitions* if there exists an orbit of size $m+1$ for the technology (t, c).

Lemma B.5. *Fix any technology (t, c). In a given setting (observable actions or hidden actions) let $Q(S)$ be the total expected payment to all agents in the best contract that induces S to exert effort.*

Consider any list of sets of agents $S_0 = \emptyset, S_1, S_2, \ldots, S_m = N$ such that $t(S_i) < t(S_{i+1})$ for all $i \in \{0, 1, 2, \ldots, m-1\}$. Assume that for any value some set in the list is optimal (for any value v and any set S that is not in the list, there exists i such that $u(S, v) \leqslant u(S_i, v)$).

The technology (t, c) has m transitions if and only if for every $k \in \{1, 2, \ldots, m-1\}$ it holds that

$$\frac{Q(S_{k+1}) - Q(S_k)}{t(S_{k+1}) - t(S_k)} > \frac{Q(S_k) - Q(S_{k-1})}{t(S_k) - t(S_{k-1})}.$$

Proof. Given a technology (t, c), let $u(S, v)$ be the principal's utility at value v, when optimally contracting with the set of agents S. Let $v_{i,j}$ be the value v in which the principal is indifferent between contracting with the set S_i or with the set S_j. By the definition of $v_{i,j}$, $u(S_j, v_{i,j}) = u(S_i, v_{i,j})$. That is, $t(S_i) \cdot v_{i,j} - Q(S_i) = t(S_j) \cdot v_{i,j} - Q(S_j)$, or equivalently $v_{i,j} = \frac{Q(S_j) - Q(S_i)}{t(S_j) - t(S_i)}$.

The claim is a result of the following two lemmas.

Lemma B.6. *For every $k \in \{1, 2, \ldots, m-1\}$*

$$u(S_k, v_{k-1,k}) > u(S_{k+1}, v_{k-1,k}) \quad \Leftrightarrow \quad \frac{Q(S_{k+1}) - Q(S_k)}{t(S_{k+1}) - t(S_k)} > \frac{Q(S_k) - Q(S_{k-1})}{t(S_k) - t(S_{k-1})}.$$

Proof. For every $k \in \{1, 2, \dots, m-1\}$

$$u(S_k, v_{k-1,k}) > u(S_{k+1}, v_{k-1,k})$$
$$\Leftrightarrow \quad t(S_k) \cdot v_{k-1,k} - Q(S_k) > t(S_{k+1}) \cdot v_{k-1,k} - Q(S_{k+1}).$$

As $v_{k-1,k} = \frac{Q(S_k)-Q(S_{k-1})}{t(S_k)-t(S_{k-1})}$, the above happens if and only if

$$Q(S_{k+1}) - Q(S_k) > \big(t(S_{k+1}) - t(S_k)\big) \cdot v_{k-1,k} = \big(t(S_{k+1}) - t(S_k)\big) \cdot \frac{Q(S_k) - Q(S_{k-1})}{t(S_k) - t(S_{k-1})}$$
$$\Leftrightarrow \quad \frac{Q(S_{k+1}) - Q(S_k)}{t(S_{k+1}) - t(S_k)} > \frac{Q(S_k) - Q(S_{k-1})}{t(S_k) - t(S_{k-1})}$$

which is what we wanted to prove. □

Lemma B.7. *Consider any list of sets of agents $S_0 = \emptyset, S_1, S_2, \dots, S_m = N$ such that $t(S_i) < t(S_{i+1})$ for all $i \in \{0, 1, 2, \dots, m-1\}$. Assume that for any value some set in the list is optimal (for any value v and any set S that is not in the list, there exists i such that $u(S, v) \leqslant u(S_i, v)$).*

The list of sets of agents $S_0, S_1, S_2, \dots, S_m$ is an orbit of size $m+1$ of the technology (t, c) (thus the technology (t, c) has m transitions) if and only if for every $k \in \{1, 2, \dots, m-1\}$ it holds that $u(S_k, v_{k-1,k}) > u(S_{k+1}, v_{k-1,k})$.

Proof. *The "if" case*: Assume that $u(S_k, v_{k-1,k}) > u(S_{k+1}, v_{k-1,k})$ for every $k \in \{1, 2, \dots, m-1\}$. To prove the claim we need to show that for every $k \in \{0, 1, 2, \dots, m\}$ there exists a value v_k for which $u(S_k, v_k) > u(S_j, v_k)$ for every $j \neq k$. Note that for any technology such v_0 and v_m clearly exist (for value close enough to 0 it is clear that contracting with no agent is the only optimal contract, and for high enough value it is clear that contracting with all agents is the only optimal contract). It remains to show that such v_k exists for every $k \in \{1, 2, \dots, m-1\}$.

We first claim that $v_{k,k+1} > v_{k-1,k}$ for any $k \in \{1, 2, \dots, m-1\}$. Assume on the contrary that $v_{k,k+1} \leqslant v_{k-1,k}$ for some $k \in \{1, 2, \dots, m-1\}$. In this case for every $v \geqslant v_{k,k+1}$ it holds that $u(S_{k+1}, v) \geqslant u(S_k, v)$; in particular for $v_{k-1,k} \geqslant v_{k,k+1}$ it holds that $u(S_{k+1}, v_{k-1,k}) \geqslant u(S_k, v_{k-1,k})$, a contradiction to the assumption that $u(S_k, v_{k-1,k}) > u(S_{k+1}, v_{k-1,k})$.

To complete the proof we show that $v_{k,k+1} > v_{k-1,k}$ for any $k \in \{1, 2, \dots, m-1\}$ implies that any $v_k \in (v_{k-1,k}, v_{k,k+1})$ satisfies that $u(S_k, v_k) > u(S_j, v_k)$ for every $j \neq k$.

We first show by induction on k that for any value $v > v_{k-1,k}$ and $j < k$, contracting with S_k gives the principal higher utility than contracting with S_j, that is, $u(S_k, v) > u(S_j, v)$ for any $j < k$ and $v > v_{k-1,k}$. Clearly the claim holds for $k = 1$ by linearity and the definition of $v_{0,1}$. Assume that we have proven the claim up to $k-1$, which means that for any $v > v_{k-2,k-1}$ and $j < k-1$ it holds that $u(S_{k-1}, v) > u(S_j, v)$. By assumption $v_{k-1,k} > v_{k-2,k-1}$ so to complete the induction step we need to show that for any value $v > v_{k-1,k}$ and $j < k$ it holds that $u(S_k, v) > u(S_j, v)$. By linearity and the definition of $v_{k-1,k}$, for any value $v > v_{k-1,k}$ it holds that $u(S_k, v) > u(S_{k-1}, v)$. By the induction hypothesis for any $v > v_{k-2,k-1}$ and $j < k-1$ it holds that $u(S_{k-1}, v) > u(S_j, v)$; thus for any $v > v_{k-1,k} > v_{k-2,k-1}$ and $j < k-1$ it holds that $u(S_{k-1}, v) > u(S_j, v)$. We conclude that for any value $v > v_{k-1,k}$ and $j < k$ it holds that $u(S_k, v) > u(S_{k-1}, v) \geqslant u(S_j, v)$.

A similar induction argument on $n-k$ shows that at any value $v < v_{k,k+1}$ and $j > k$ it holds that $u(S_k, v) > u(S_j, v)$. Combining the two claims we derive that for any value $v \in (v_{k-1,k}, v_{k,k+1})$ and $j \neq q$ it holds that $u(S_k, v) > u(S_j, v)$, as needed.

The "only if" case: Assume that for some $k \in \{1, 2, \ldots, m-1\}$ it holds that $u(S_k, v_{k-1,k}) \leqslant u(S_{k+1}, v_{k-1,k})$. As $t(S_{k+1}) > t(S_k)$ and for any set T the utility $u(T, v)$ is linear in v with slope $t(T)$, we conclude that for every $v > v_{k-1,k}$ it holds that $u(S_k, v) < u(S_{k+1}, v)$. On the other hand, for every $v \leqslant v_{k-1,k}$ it holds that $u(S_k, v) \leqslant u(S_{k-1}, v)$. We conclude that there does not exist a value v_k for which $u(S_k, v_k) > u(S_j, v_k)$ for every $j \neq k$, which contradicts the second condition in the definition of an orbit that includes S_k. □

This completes the proof of Lemma B.5. □

We can now use Lemma B.5 to derive Theorems 4.6 and 4.10.

For any $k \in \{0, 1, \ldots, n\}$, let S_k^* be a set of size k such that $t(S_k^*) \geqslant t(S_k)$ for any set S_k of size k. Observe that $S_0^* = \emptyset, S_1^*, S_2^*, \ldots, S_n^* = N$ satisfy $t(S_i^*) < t(S_{i+1}^*)$ for all $i \in \{0, 1, 2, \ldots, m-1\}$.

In the observable-actions case, for any value some set in the list is optimal, as for any value v and any set S that is not in the list, if $|S| = k$ then $u(S, v) \leqslant u(S_k^*, v)$. Now, for general technologies in the observable-actions case, Theorem 4.6 which presents a necessary and sufficient condition for n transitions is a direct result of Lemma B.5 and the following observation.

Observation B.8. *Consider a technology (t, c) with identical costs and consider the observable-actions case in which for a set S the payment is $Q(S) = c|S|$.*

A technology success function t exhibits (strictly) decreasing returns to maximal scale (DRMS) if and only if for every $k \in \{1, 2, \ldots, n-1\}$ it holds that

$$\frac{Q(S_{k+1}^*) - Q(S_k^*)}{t(S_{k+1}^*) - t(S_k^*)} > \frac{Q(S_k^*) - Q(S_{k-1}^*)}{t(S_k^*) - t(S_{k-1}^*)}.$$

Proof. Recall that a technology success function t exhibits (strictly) decreasing returns to maximal scale (DRMS) if for every $k \in \{1, 2, \ldots, n-1\}$ it holds that

$$t\big(S_k^*\big) - t\big(S_{k-1}^*\big) > t\big(S_{k+1}^*\big) - t\big(S_k^*\big).$$

The claim follows from the fact that in the observable-actions case it holds that $Q(S_{k+1}^*) - Q(S_k^*) = Q(S_k^*) - Q(S_{k-1}^*) = c$. □

For anonymous technologies in the hidden-actions case, the condition for n transitions (Theorem 4.10) is a direct result of Lemma B.5 and the fact that (strictly) increasing relative marginal payment ($IRMP_{anon}$) is exactly the needed condition for Lemma B.5 to hold (the other conditions needed for Lemma B.5 hold trivially as $t_{k+1} > t_k$ for every $k \in \{0, 1, \ldots, n-1\}$ and clearly for an anonymous technology all sets of a given size gives the same utility).

B.3. Relations between the properties

We have presented various properties that determine the number of transitions points in the different cases (observable and hidden actions, general and anonymous technologies). What are the relations between the different properties? Does any of them imply any other? In addition, does any of them relate to the two standard properties of IRS and DRS? We consider these questions next. To maintain the flow of this section, some technical proofs are deferred to Appendix B.3.1.

We first consider properties that relate to a single transition: *OP*, *UPC*, and *IRS*. We start by observing that any technology that exhibits IRS also exhibits UPC.

Lemma B.9. *Any technology (t, c) that exhibits IRS also exhibits UPC.*

Proof. Assume on the contrary that the technology exhibits IRS and not UPC. If the technology does not exhibit UPC then there exists a non-empty set $S \subsetneq N$ such that

$$\frac{|S|}{|N|} \leqslant \frac{t(S) - t(\emptyset)}{t(N) - t(\emptyset)}.$$

Denote $|S| = k$; note that $k > 0$ and $k < n$. Fix an order on the agents such that all agents in S come before any agents that are not in S. That is, the order is $j_1, j_2, \ldots, j_n$ such that $j_1, j_2, \ldots, j_k$ are in S. Let S_i denote the set $S_i = \{j_1, j_2, \ldots, j_i\}$.

As the technology exhibits IRS, it holds that for any $i \in \{2, \ldots, k\}$, $t(S_i) - t(S_{i-1}) > t(S_{i-1}) - t(S_{i-2})$, and thus for any $i \in \{1, \ldots, k-1\}$ it holds that $t(S_k) - t(S_{k-1}) > t(S_i) - t(S_{i-1})$. Therefore by summation

$$\begin{aligned}(k-1)\big(t(S_k) - t(S_{k-1})\big) &= \sum_{i=1}^{k-1}\big(t(S_k) - t(S_{k-1})\big) > \sum_{i=1}^{k-1}\big(t(S_i) - t(S_{i-1})\big) \\ &= t(S_{k-1}) - t(S_0).\end{aligned}$$

Thus by adding $t(S_k) - t(S_{k-1})$ we derive that $t(S_k) - t(S_{k-1}) > \frac{t(S_k)-t(S_0)}{k} = \frac{t(S)-t(\emptyset)}{k}$. Recall that we assume in contradiction that $\frac{k}{n} \leqslant \frac{t(S)-t(\emptyset)}{t(N)-t(\emptyset)}$, thus $t(S_k) - t(S_{k-1}) > \frac{t(N)-t(\emptyset)}{n}$.

As for any $i \in \{(k+1), \ldots, n\}$, by IRS it holds that $t(S_k) - t(S_{k-1}) < t(S_i) - t(S_{i-1})$,

$$\begin{aligned}(n-k)\big(t(S_k) - t(S_{k-1})\big) &= \sum_{i=k+1}^{n}\big(t(S_k) - t(S_{k-1})\big) < \sum_{i=k+1}^{n}\big(t(S_i) - t(S_{i-1})\big) \\ &= t(N) - t(S_k)\end{aligned}$$

which implies that $t(N) - t(S) > \frac{n-k}{n}(t(N) - t(\emptyset))$. With the assumption that $\frac{k}{n}(t(N) - t(\emptyset)) \leqslant t(S) - t(\emptyset)$ we conclude that $t(N) - t(\emptyset) = (t(N) - t(S)) + (t(S) - t(\emptyset)) > t(N) - t(\emptyset)$, a contradiction. $\square$

Thus, IRS is sufficient to ensure a single transition in the observable-actions case.

We next show that even for anonymous technologies none of the other five possible implications is true.

Proposition B.10. *Even for anonymous technologies*:

- OP_{anon} *does not imply* IRS_{anon} *or* UPC_{anon}.
- UPC_{anon} *does not imply* IRS_{anon} *or* OP_{anon}.[22]
- IRS_{anon} *does not imply* OP_{anon}.

[22] It is easy to show that for the special case of $n = 2$ the UPC_{anon} and IRS_{anon} conditions are equivalent, but UPC_{anon} does not imply IRS_{anon} in general.

From the fact that UPC_{anon} does not imply OP_{anon} and vice versa, we conclude that it is not necessarily the case that one transition in the observable-actions case implies one transition in the hidden-actions case, or vice versa.

Although Proposition B.10 shows that *IRS* is neither necessary nor sufficient to get *OP* (and thus one transition in the hidden-actions case) we next show that even for general technologies, a slight strengthening of *IRS* implies *OP*.

Definition B.11. A technology success function t exhibits **strong increasing returns to scale (strong IRS)** if for every i and every $a_{-i} < b_{-i}$ it holds that $\frac{\Delta_i(a_{-i})}{t(1,a_{-i})} \leqslant \frac{\Delta_i(b_{-i})}{t(1,b_{-i})}$.

Note that this is a stronger condition than *IRS*. *IRS* only requires that $\Delta_i(a_{-i}) < \Delta_i(b_{-i})$ while the new condition requires that $\Delta_i(a_{-i}) \cdot t(1, b_{-i}) \leqslant \Delta_i(b_{-i}) \cdot t(1, a_{-i})$, and as $a_{-i} < b_{-i}$ it holds that $t(1, a_{-i}) < t(1, b_{-i})$ and thus the strong IRS condition is harder to satisfy.

Lemma B.12. *Any technology (t, c) that exhibits strong IRS exhibits OP.*

Proof. As we observed strong *IRS* implies *IRS*. By Lemma B.9 any technology (t, c) that exhibits *IRS* also exhibits *UPC*. Thus it is sufficient to show that if a technology exhibits strong *IRS* then for any non-empty set $S \subsetneq N$ it holds that $\frac{Q(S)}{Q(N)} \geqslant \frac{|S|}{|N|}$. Strong *IRS* implies that for every such S that contains i it holds that $\frac{t(S)}{t(S)-t(S\setminus\{i\})} \geqslant \frac{t(N)}{t(N)-t(N\setminus\{i\})}$. By averaging over all agents we get

$$\frac{1}{|S|}\sum_{i\in S}\frac{t(S)}{t(S)-t(S\setminus\{i\})} \geqslant \frac{1}{|N|}\sum_{i\in N}\frac{t(N)}{t(N)-t(N\setminus\{i\})}$$

which implies that

$$\frac{Q(S)}{Q(N)} = \frac{t(S)\sum_{i\in S}\frac{c}{t(S)-t(S\setminus\{i\})}}{t(N)\sum_{i\in N}\frac{c}{t(N)-t(N\setminus\{i\})}} \geqslant \frac{|S|}{|N|}$$

as needed. □

We next consider properties that relate to n transitions: *DRMS*, $IRMP_{anon}$ and *DRS*. We first consider the two properties that are defined for general technologies, *DRMS* and *DRS*. We show that *DRMS* does not imply *DRS* or vice versa.

Proposition B.13. *DRMS does not imply DRS and DRS does not imply DRMS.*

Finally we consider the relations between the properties for anonymous technologies. As pointed out before, for anonymous technologies, *DRMS* is equivalent to DRS_{anon}; thus we only need to consider the relation between $IRMP_{anon}$ and DRS_{anon}. We prove that neither implies the other.[23]

Proposition B.14. *For anonymous technologies, neither $IRMP_{anon}$ implies DRS_{anon} nor DRS_{anon} implies $IRMP_{anon}$.*

[23] It is easy to show that for the special case of $n = 2$ the two conditions are equivalent, but not in general.

We conclude that for anonymous technologies it is not necessarily the case that n transitions in the observable-actions case imply n transitions in the hidden-actions case, or vice versa.

B.3.1. The proofs

Proposition B.10. *Even for anonymous technologies*:

- OP_{anon} *does not imply* IRS_{anon} *or* UPC_{anon}.
- UPC_{anon} *does not imply* IRS_{anon} *or* OP_{anon}.
- IRS_{anon} *does not imply* OP_{anon}.

Proof. To observe that OP_{anon} does not imply UPC_{anon} or IRS_{anon}, consider the anonymous technology with three agents defined as $t_0 = 0.08$, $t_1 = 0.1$, $t_2 = 0.11$, $t_3 = 0.14$. The technology exhibits OP_{anon} as $Q_1 = 5$, $Q_2 = 22$, and $Q_3 = 14$, and $Q_1/Q_3 = 5/14 > (t_1 - t_0)/(t_3 - t_0) = 1/3$ and $Q_2/Q_3 = 22/14 > (t_2 - t_0)/(t_3 - t_0) = 1/2$. The technology does not exhibit UPC_{anon} as $1/3 \leqslant (t_1 - t_0)/(t_3 - t_0) = 1/3$, nor IRS_{anon} as $t_2 - t_1 = 0.01 < t_1 - t_0 = 0.02$.

To show that UPC_{anon} does not imply IRS_{anon}, consider the anonymous technology with three agents defined as $t_0 = 1/10$, $t_1 = 3/10$, $t_2 = 45/100$, $t_3 = 1$. This technology exhibits UPC_{anon} as $1/3 > (t_1 - t_0)/(t_3 - t_0) = (3/10 - 1/10)/(1 - 1/10) = 2/9$ and $2/3 > (t_2 - t_0)/(t_3 - t_0) = (45/100 - 1/10)/(1 - 1/10) = 7/18$. It does not exhibit IRS_{anon} as $t_2 - t_1 = 45/100 - 3/10 = 3/20 < 1/5 = 3/10 - 1/10 = t_1 - t_0$.

To show that neither UPC_{anon} nor IRS_{anon} implies OP_{anon}, consider the anonymous technology with two agents defined as $t_0 = 0.1$, $t_1 = 0.4$, $t_2 = 0.9$. This technology exhibits UPC_{anon} as $1/2 > (t_1 - t_0)/(t_2 - t_0) = 3/8$, and it exhibits IRS_{anon} as $t_2 - t_1 = 1/2 > 3/10 = t_1 - t_0$. This technology does not exhibit OP_{anon} as $Q_1 = 4/3$ and $Q_2 = 18/5$, so $Q_1/Q_2 = 10/27 < (t_1 - t_0)/(t_2 - t_0) = 3/8$. $\square$

Proposition B.13. *DRMS does not imply DRS and DRS does not imply DRMS.*

Proof. To see that *DRMS* does not imply *DRS*, consider the anonymous technology defined as $t(\emptyset) = 1/10$, $t(\{1\}) = 3/10$, $t(\{2\}) = 9/10$ and $t(\{1,2\}) = 1$. Clearly $S_2^* = \{1,2\}$, $S_1^* = \{2\}$ and $S_0^* = \emptyset$, and DRMS holds as $t(S_1^*) - t(S_0^*) = 8/10 > 1/10 = t(S_2^*) - t(S_1^*)$. Yet this technology does not exhibit DRS as $t(\{1,2\}) - t(\{2\}) = 1/10 > 3/10 = t(\{1\}) - t(\emptyset)$.

While for anonymous technologies it is clear that *DRS* implies *DRMS*, we next present a technology that is non-anonymous and satisfies *DRS* but not *DRMS*, showing that *DRS* does not imply *DRMS* in general. The technology has four agents, named A, B, C, and D, and is defined as follows. $t(A) = t(B) = t(C) = 8/100 = 2/25$, $t(D) = 1/10$, $t(AB) = t(BC) = t(CD) = t(AC) = t(AD) = t(BD) = 15/100 = 3/20$, $t(ABC) = 21/100$, $t(ABD) = t(BCD) = t(ACD) = 19/100$ and $t(ABCD) = 22/100$. It holds that $t(S_1^*) = t(D) = 1/10$, $t(S_2^*) = 15/100$, $t(S_3^*) = t(ABC) = 21/100$. *DRMS* is violated as $5/100 = t(S_2^*) - t(S_1^*) < t(S_3^*) - t(S_2^*) = 6/100$. It is easy to verify that the technology satisfies *DRS*. $\square$

Proposition B.14. *For anonymous technologies, neither* $IRMP_{anon}$ *implies* DRS_{anon} *nor* DRS_{anon} *implies* $IRMP_{anon}$.

Proof. We first show that $IRMP_{anon}$ does not imply DRS_{anon}. Consider the anonymous technology with two agents defined as $t_0 = 0.1$, $t_1 = 0.4$, $t_2 = 0.9$. This technology exhibits $IRMP_{anon}$

as $\frac{Q_2-Q_1}{t_2-t_1} = \frac{18/5-4/3}{9/10-4/10} = 68/15 > 40/9 = \frac{4/3-0}{4/10-1/10} = \frac{Q_1-Q_0}{t_1-t_0}$, but it does not exhibit DRS_{anon} as $t_2 - t_1 = 1/2 > 3/10 = t_1 - t_0$.

To show that DRS_{anon} does not imply $IRMP_{anon}$ consider the anonymous technology with four agents defined as $t_0 = 1/1000$, $t_1 = 401/1000$, $t_2 = 800/1000$, $t_3 = 901/1000$, $t_4 = 1$. This technology clearly exhibits DRS_{anon}, but it does not exhibit $IRMP_{anon}$ as $Q_2 = 2 \cdot (800/1000)/(399/1000) = 1600/399$, $Q_3 = 3 \cdot (901/1000)/(101/1000) = 2703/101$ and $Q_4 = 4 \cdot 1/(99/1000) = 4000/99$, and thus $\frac{Q_4-Q_3}{t_4-t_3} = \frac{4000/99-2703/101}{1-901/1000} < 138 < 225 < \frac{2703/101-1600/399}{901/1000-800/1000} = \frac{Q_3-Q_2}{t_3-t_2}$. □

B.4. *Properties of anonymous* AND *and* OR *technologies*

In this section we present the properties of anonymous AND and OR that are used to show Observations 4.11 and 4.12.

We first show that any AND technology (even non-anonymous) exhibits IRS.

Observation B.15. *Any AND technology exhibits IRS.*

Proof. Assume that agent i's effort a_i results in success probability of $r_i^a > 0$ in his task. For an AND technology the technology success probability is $t(r^a) = \prod_{i \in N} r_i^a$, where $r^a = (r_1^a, r_2^a, \dots, r_n^a)$. Fix some agent i and two profiles of actions a and b such that $a_{-i} < b_{-i}$. To prove that the technology exhibits IRS we need to show that $\Delta_i(a_{-i}) < \Delta_i(b_{-i})$. Let $r^a, r^b \in [0, 1]^n$ be the two vectors of individual success probabilities in the agent's tasks that correspond to action vectors a and b, respectively. For vectors a and b such that $a_{-i} < b_{-i}$ it holds that for any j, $r_j^b \geqslant r_j^a > 0$, and the inequality is strict at least for one j. Thus $\prod_{j \neq i} r_j^b > \prod_{j \neq i} r_j^a$. We conclude that $\Delta_i(b_{-i}) = t(r_i^b, r_{-i}^b) - t(r_i^a, r_{-i}^b) = (r_i^b - r_i^a) \cdot \prod_{j \neq i} r_j^b > (r_i^b - r_i^a) \cdot \prod_{j \neq i} r_j^a = t(r_i^b, r_{-i}^a) - t(r_i^a, r_{-i}^a) = \Delta_i(a_{-i})$, as we needed to show. □

The lemma combined with Corollary 4.7 implies Observation 4.11.

Lemma B.16. *An anonymous AND technology exhibits both under-proportional contribution (UPC_{anon}) and over-payment (OP_{anon}).*

Proof. By Observation B.15 an anonymous AND technology exhibits IRS, and by Lemma B.9 this implies that it exhibits UPC_{anon}. Next we show that it exhibits over-payment (OP_{anon}). Let t be an anonymous AND technology. Observe that $t_k = \delta^k \gamma^{n-k}$. As for any k, $\frac{t_k}{t_k - t_{k-1}} = \frac{\delta}{\delta - \gamma}$, it implies that $\frac{Q_k}{Q_n} = \frac{(\frac{c \cdot k \cdot t_k}{t_k - t_{k-1}})}{(\frac{c \cdot n \cdot t_n}{t_n - t_{n-1}})} = \frac{k}{n}$. Thus, OP is equivalent to UPC for an anonymous AND technology, and as we have seen any such technology exhibits an under-proportional contribution (UPC_{anon}). □

The next lemma combined with Theorem 4.10 and Corollary 4.8 implies Observation 4.12.

Lemma B.17. *Any anonymous OR technology exhibits DRS and $IRMP_{anon}$, and thus has all n transitions in both the hidden-actions and observable-actions cases.*

Proof. Fix any anonymous OR technology. Let t_k denote the success probability when k out of n agents exert effort. It holds that $t_k = 1-(1-\delta)^k(1-\gamma)^{n-k} = 1-r^k(1-\gamma)^n$ for $r = \frac{1-\delta}{1-\gamma}$, and $\Delta_k = t_k - t_{k-1} = (1-r^k(1-\gamma)^n) - (1-r^{k-1}(1-\gamma)^n) = r^{k-1}(1-\gamma)^n(\delta-\gamma)$. This implies that $\Delta_{k+1} = r\Delta_k$, and as $r<1$ we conclude that OR exhibits DRS. Next we show that it also exhibits $IRMP_{anon}$.

To show that the technology exhibits $IRMP_{anon}$ we need to show that for any $k \in \{1, 2, \ldots, n-1\}$,

$$\frac{Q_{k+1}-Q_k}{\Delta_{k+1}} > \frac{Q_k - Q_{k-1}}{\Delta_k}$$

where $Q_k = \frac{c \cdot k \cdot t_k}{\Delta_k}$ is the total expected payment in the best contract for which there exists an equilibrium with k agents exerting effort.

$$\begin{aligned}\frac{Q_{k+1}-Q_k}{\Delta_{k+1}} &= \frac{c}{\Delta_{k+1}} \cdot \left(\frac{(k+1)\cdot t_{k+1}}{\Delta_{k+1}} - \frac{k \cdot t_k}{\Delta_k}\right) \\ &= \frac{c((k+1)\cdot t_{k+1} - r \cdot k \cdot t_k)}{(\Delta_{k+1})^2} = \frac{c(r \cdot t_k + (1-r)\cdot(k+1))}{(\Delta_{k+1})^2}\end{aligned}$$

where the last equality is derived from $t_{k+1} = r \cdot t_k + 1 - r$.

We use the facts that $\Delta_{k+1} = r \cdot \Delta_k$, and that $r \cdot t_{k-1} = t_k - (1-r)$ to conclude:

$$\begin{aligned}&\frac{Q_{k+1}-Q_k}{\Delta_{k+1}} > \frac{Q_k - Q_{k-1}}{\Delta_k} \\ \Leftrightarrow\quad & \frac{c(r\cdot t_k + (1-r)\cdot(k+1))}{(\Delta_{k+1})^2} > \frac{c(r \cdot t_{k-1} + (1-r)\cdot k)}{(\Delta_k)^2} \\ \Leftrightarrow\quad & r \cdot t_k + (1-r)\cdot(k+1) > r^2 \cdot \big(r \cdot t_{k-1} + (1-r)\cdot k\big) \\ \Leftrightarrow\quad & r \cdot t_k + (1-r)\cdot(k+1) > r^2 \cdot \big(t_k + (1-r)\cdot(k-1)\big) \\ \Leftrightarrow\quad & r(1-r)t_k > (1-r)\big(r^2(k-1) - (k+1)\big) \\ \Leftrightarrow\quad & r \cdot t_k > r^2(k-1) - (k+1).\end{aligned}$$

And this holds as $0 < r < 1$; thus $r \cdot t_k > 0$ while $r^2(k-1) - (k+1) < 0$. □

Appendix C. Agent selection

Observation 5.2. *Fix a technology (t, c). In the observable-actions case if agent j is more productive than agent i then it holds that for every v and for every optimal contract S^*, if $i \in S^*$ then $j \in S^*$.*

Proof. Assume on the contrary that for some value v it holds that S^* is an optimal contract and $i \in S^*$ while $j \notin S^*$. We show that $S' = S^* \setminus \{i\} \cup \{j\}$ is a better contract at v, which contradicts the optimality of S^*. As we assume that the cost of effort is the same for all agents, the two contracts have the same cost. Yet, as j has higher productivity, S' provides higher success probability than S^*. We conclude that the principal's utility is higher with S' than with S^*, a contradiction. □

Theorem 5.5. *Fix a technology (t, c). In the hidden-actions case if agent j is more productive than agent i and also a better productivity initiator than i then it holds that for every v and for every optimal contract S^*, if $i \in S^*$ then $j \in S^*$.*

Proof. Assume that for some value v it holds that S^* is an optimal contract and $i \in S^*$ while $j \notin S^*$. We show that $S' = S^* \setminus \{i\} \cup \{j\}$ is a better contract at v, by showing that it ensures higher success probability at a lower payment. Higher success probability is immediate: as j is more productive than i it holds that $t(S') > t(S^*)$.

Higher productivity ensures that agent j is paid less with contract S', than agent i is paid with contract S^*. Productivity initiation ensures that each agent $k \in S^* \setminus \{i\}$ is paid less with the contract S' than he is paid with the contract S^*. Thus the utility of the principal is higher with S', a contradiction. □

Appendix D. Price of unobservability

Theorem 6.1. *For any technology (t, c) with n agents that exhibits DRS, it holds that $POU_S \leqslant POU_P \leqslant 2n$.*

Proof. By Observation 2.8, for any technology it holds that $POU_S \leqslant POU_P$; thus it is sufficient to show that $POU_P \leqslant 2n$.

Let $i^* = argmax_i\{t(\{i\})\colon\ i \in N\}$; that is, i^* is some agent that maximizes the probability of success among all sets of size 1. For simplicity, we slightly abuse notation and write $t(i^*)$ to mean $t(\{i^*\})$.

Let v^*_{oa} and v^* denote the values for which the principal is indifferent between contracting with 0 agents or with agent i^* in the observable-actions case and the hidden-actions case, respectively. Simple calculation reveals that $v^* = \frac{ct(i^*)}{\Delta^2_{i^*}}$, where $\Delta_{i^*} = t(i^*) - t(\emptyset)$.

We first show that for every technology that exhibits DRS, for every $S \subseteq N$, it holds that

$$t(S) - t(\emptyset) \leqslant |S|\big(t\big(i^*\big) - t(\emptyset)\big). \tag{D.1}$$

Let $S = \{i_1, \ldots, i_{|S|}\}$, and let $S^{\leqslant j} = \{i_1, \ldots, i_j\}$ for every $1 \leqslant j \leqslant |S|$. For every S, it holds that $t(S) - t(\emptyset) = \sum_{j=1}^{|S|}(t(S^{\leqslant j}) - t(S^{\leqslant (j-1)}))$. Since t exhibits DRS, $t(S^{\leqslant j}) - t(S^{\leqslant (j-1)}) \leqslant t(\{i_j\}) - t(\emptyset) \leqslant t(i^*) - t(\emptyset)$ for every j. Therefore, $\sum_{j=1}^{|S|}(t(S^{\leqslant j}) - t(S^{\leqslant (j-1)})) \leqslant |S|(t(i^*) - t(\emptyset))$, which establishes Eq. (D.1).

Let $v^{oa}_{S,S'}$ denote the value v in which the principal is indifferent between contracting with the sets S and S' in the observable-actions case. Similarly, let $v^{ha}_{S,S'}$ denote the indifference value between S and S' in the hidden-actions case.

The following lemma will be used in the proof.

Lemma D.1. *For every technology that exhibits DRS, the empty set is an optimal contract for every $v < v^{oa}_{\emptyset,\{i^*\}}$ both in the observable- and the hidden-actions cases.*

Proof. One can easily verify that to prove the lemma it is sufficient to show that $v_{\emptyset,\{i^*\}} \leqslant v_{\emptyset,S}$ for every set $S \subseteq N$. We show that this inequality holds in both the observable- and the hidden-actions cases.

In the observable-actions case, for every $S \subseteq N$ it holds that $v^{oa}_{\emptyset,S} = \frac{c|S|}{t(S)-t(\emptyset)}$. Therefore, the claim holds if and only if $\frac{c|S|}{t(S)-t(\emptyset)} \geqslant \frac{c}{t(i^*)-t(\emptyset)}$. By rearrangement we get $t(S) - t(\emptyset) \leqslant |S|(t(i^*) - t(\emptyset))$. But the last inequality holds for every DRS technology by Eq. (D.1).

We conclude that in the observable-actions case for any $v < v^{oa}_{\emptyset,\{i^*\}}$ the utility of the principal from the empty set is at least as high as his utility from any other set S. The utility from the empty set is the same both in the hidden-actions and the observable-actions cases, while the utility from any other set S is smaller in the hidden-actions case (as the expected payment is larger); thus for any $v < v^{oa}_{\emptyset,\{i^*\}}$ the empty set is also optimal in the hidden-actions case. □

In the remainder of this proof, we denote the optimal contract at v in the observable-actions case by S_v. We show that the ratio between the principal's optimal utility in the observable- and the hidden-actions cases is bounded by $2n$ for every v.

By Lemma D.1 for every $v < v^*_{oa}$ the empty contract is optimal in both the hidden- and the observable-actions cases. Thus, the utility ratio for this interval is 1.

We next consider the interval $v^*_{oa} \leqslant v \leqslant v^*$. Clearly, for such values, it holds that $|S_v| \geqslant 1$. Since the principal's utility in the hidden-actions case is never less than $vt(\emptyset)$, the utility ratio is at most $\frac{t(S_v)v-c|S_v|}{vt(\emptyset)}$, which is at most $\frac{t(S_v)}{t(\emptyset)} - \frac{c|S_v|}{v^*t(\emptyset)}$ for $v \leqslant v^*$. It remains to show that

$$\frac{t(S_v)}{t(\emptyset)} - \frac{c|S_v|}{v^*t(\emptyset)} \leqslant 2n. \tag{D.2}$$

Substituting $v^* = \frac{c\cdot t(i^*)}{\Delta^2_{i^*}}$, we get that

$$\begin{aligned}\frac{t(S_v)}{t(\emptyset)} - \frac{c|S_v|}{v^*t(\emptyset)} &= \frac{t(S_v)}{t(\emptyset)} - \frac{c|S_v|\Delta^2_{i^*}}{c\cdot t(i^*)t(\emptyset)}\\ &= \frac{t(i^*)t(S_v) - |S_v|(t(i^*)^2 - 2t(\emptyset)t(i^*) + t(\emptyset)^2)}{t(i^*)t(\emptyset)}\\ &= 2|S_v| + \frac{t(S_v) - |S_v|t(i^*)}{t(\emptyset)} - \frac{|S_v|t(\emptyset)}{t(i^*)}.\end{aligned}$$

We next establish that the last expression is bounded from above by $2|S_v|$ for every DRS technology, which is in turn bounded by $2n$. It is sufficient to show that $t(S_v) - |S_v|t(i^*) \leqslant 0$. But this follows from Eq. (D.1) as $t(S_v) - |S_v|t(i^*) \leqslant t(\emptyset)(1 - |S_v|) \leqslant 0$, as required.

Finally, we establish the bound for values $v \geqslant v^*$. The principal's utility in the hidden-actions case for $v \geqslant v^*$ is at least $t(i^*)(v - c/\Delta_{i^*})$. Therefore, the utility ratio is at most $\frac{t(S_v)v-c|S_v|}{t(i^*)(v-c/\Delta_{i^*})}$. Clearly, for $v \geqslant v^*$, it holds that $v - c/\Delta_{i^*} > 0$. We consider two cases.

Case (a). $t(S_v) \geqslant |S_v|(t(i^*) - t(\emptyset))$. We get

$$\frac{t(S_v)v - c|S_v|}{t(i^*)(v - c/\Delta_{i^*})} \leqslant \frac{t(S_v)v^* - c|S_v|}{t(i^*)(v^* - c/\Delta_{i^*})} = \frac{t(S_v)v^* - c|S_v|}{t(\emptyset)v^*} \leqslant 2n,$$

where the first inequality follows from $v \geqslant v^*$ and the assumption $t(S_v) \geqslant |S_v|(t(i^*) - t(\emptyset))$, the equality follows from the definition of v^* (which implies $t(i^*)(v^* - c/\Delta_{i^*}) = t(\emptyset)v^*$), and the last inequality follows from Eq. (D.2).

Case (b). $t(S_v) < |S_v|(t(i^*) - t(\emptyset))$. We get

$$\frac{t(S_v)v - c|S_v|}{t(i^*)(v - c/\Delta_{i^*})} = \frac{t(S_v)}{t(i^*)}\left(1 - \frac{c(\frac{|S_v|}{t(S_v)} - \frac{1}{\Delta_{i^*}})}{v - \frac{c}{\Delta_{i^*}}}\right) < \frac{t(S_v)}{t(i^*)}.$$

However, by Eq. (D.1), $t(S_v) \leqslant |S_v|t(i^*) - t(\emptyset)(|S_v| - 1) \leqslant |S_v|t(i^*)$, and thus the last expression is bounded from above by $|S_v| \leqslant n$. This establishes the assertion of the theorem. □

Theorem 6.2. *There exists an anonymous technology with n agents that exhibits DRS for which* $POU_P \geqslant n/2$.

Proof. Fix $0 < t_0 < \frac{1}{2n}$, and consider the technology in which $t_k = t_0(k+1) - k^2\epsilon$ for a sufficiently small $\epsilon > 0$ that will be determined later on. It holds that $\Delta_k = t_k - t_{k-1} = t_0 - \epsilon(2k-1)$. It is easy to verify that this technology exhibits DRS for any $\epsilon > 0$. We claim that for a sufficiently small ϵ this technology exhibits IRMP. Recall that in order to show that our technology exhibits IRMP, we need to show that $\frac{Q_{k+1}-Q_k}{t_{k+1}-t_k} > \frac{Q_k - Q_{k-1}}{t_k - t_{k-1}}$ for all $k \geqslant 1$ such that $k < n$. In our case $Q_k = \frac{c \cdot k \cdot t_k}{\Delta_k} = \frac{c \cdot k(t_0 + t_0 k - k^2\epsilon)}{t_0 - \epsilon(2k-1)}$. It holds that $\lim_{\epsilon \to 0} \frac{Q_{k+1}-Q_k}{t_{k+1}-t_k} = \frac{2 \cdot c(k+1)}{t_0}$, and $\lim_{\epsilon \to 0} \frac{Q_k - Q_{k-1}}{t_k - t_{k-1}} = \frac{2 \cdot c \cdot k}{t_0}$. Thus, $\exists \epsilon' > 0$ s.t. $\forall \epsilon < \epsilon'$, $\frac{Q_{k+1}-Q_k}{t_{k+1}-t_k} > \frac{Q_k - Q_{k-1}}{t_k - t_{k-1}} \Leftrightarrow \frac{2 \cdot c(k+1)}{t_0} > \frac{2 \cdot c \cdot k}{t_0}$, which holds $\forall k \geqslant 1$.

Let v^* be the value for which the principal is indifferent between contracting with 0 or 1 agents in the hidden-actions case. At $v = v^*$ it holds that $t_1 \cdot (v - \frac{1}{\Delta_1}) = t_0 \cdot v$ and thus $v^* = \frac{c \cdot t_1}{\Delta_1^2} = c \cdot \frac{2t_0 - \epsilon}{(t_0 - \epsilon)^2}$. (Note that we look at v^* as a function of ϵ but to simplify notation we omit the explicit functional dependency.)

Since the technology exhibits IRMP, contracting with 0 agents is optimal at v^* in the hidden-actions case, and the principal's utility at v^* is exactly $t_0 v^*$.

In the observable-actions case, the principal's utility at v^* if contracting with k agents is $t_k v^* - c \cdot k = (t_0(k+1) - k^2\epsilon) \cdot c \cdot (\frac{2t_0 - \epsilon}{(t_0 - \epsilon)^2}) - c \cdot k$. It holds that $\lim_{\epsilon \to 0} t_k v^* - c \cdot k = c \cdot (k+2)$, which increases in k. Thus, $\exists \epsilon'' > 0$ s.t. $\forall \epsilon < \epsilon''$, it is optimal for the principal to contract with n agents, in which case the obtained utility is $t_n v^* - c \cdot n$. Thus, the ratio between the principal's utility in the observable- and the hidden-actions cases at v^* is $\frac{t_n v^* - c \cdot n}{t_0 v^*}$. It holds that $\lim_{\epsilon \to 0} \frac{t_n v^* - c \cdot n}{t_0 v^*} = \frac{n}{2} + 1$. Thus, $\exists \epsilon''' > 0$ s.t. $\forall \epsilon < \epsilon'''$, $\frac{t_n v^* - c \cdot n}{t_0 v^*} \geqslant \frac{n}{2}$. We conclude that $\forall \epsilon < \min(\epsilon', \epsilon'', \epsilon''')$, it holds that $POU_p \geqslant \frac{n}{2}$. □

Lemma 6.4. *For any anonymous technology* (t, c) *that exhibits both UPC and OP it holds that*

$$POU_S = POU_P = 1 + \frac{t_{n-1}}{t_0} - \frac{t_{n-1}}{t_n}$$

and it is obtained at the unique transition point of the hidden-actions case.

Proof. By Theorem 4.3 the technology has a single transition in both the hidden-actions and the observable-actions cases. Let v_{ha} be the value at which the transition occurs in the hidden-actions case, and let v_{oa} be the value at which the transition occurs in the observable-actions case. The transition value is the value in which the principal is indifferent between contracting with 0 agents and contracting with n agents. Thus v_{oa} solves the equation $v_{oa} \cdot t_n - c \cdot n = v_{oa} \cdot t_0$, so $v_{oa} = \frac{c \cdot n}{t_n - t_0}$. Additionally, v_{ha} solves the equation $t_n \cdot (v_{ha} - \frac{c \cdot n}{t_n - t_{n-1}}) = v_{ha} \cdot t_0$, so $v_{ha} = \frac{c \cdot n}{t_n - t_0} \cdot \frac{t_n}{t_n - t_{n-1}} = v_{oa} \cdot \frac{t_n}{t_n - t_{n-1}}$.

As we assumed that $t_0 > 0$ it follows that $t_{n-1} > 0$. Thus, $\frac{t_n}{t_n - t_{n-1}} > 1$, and therefore $v_{ha} > v_{oa}$ (i.e., the transition in the hidden-actions case occurs at a larger value than in the observable-actions case). By Lemma 2.9, both the social POU and the principal's POU are obtained at one of the two transition points (v_{oa} and v_{ha}). As at v_{oa} no agent is contracted in the hidden-actions case, both the social welfare ratio and the principal's utility ratio are 1 at v_{oa}. Thus, we only need to check the ratios at v_{ha}.

As v_{ha} is a transition point of the hidden-actions case, between 0 and n, the worst social welfare at this point is obtained with 0 agents exerting effort, and it is $v_{ha} \cdot t_0$. This is also the principal's utility at v_{ha}. Thus, the social and principal's POU are $POU_S = POU_P = \frac{v_{ha} \cdot t_n - c \cdot n}{v_{ha} \cdot t_0} = \frac{t_n}{t_0} - \frac{c \cdot n}{v_{ha} \cdot t_0}$. As $\frac{c \cdot n}{v_{ha} \cdot t_0} = \frac{c \cdot n}{\frac{c \cdot n}{t_n - t_0} \cdot \frac{t_n}{t_n - t_{n-1}} \cdot t_0} = (\frac{t_n}{t_0} - 1) \cdot (1 - \frac{t_{n-1}}{t_n})$, we derive that

$$\begin{aligned} POU_S = POU_P &= \frac{t_n}{t_0} - \frac{c \cdot n}{v_{ha} \cdot t_0} = \frac{t_n}{t_0} - \left(\frac{t_n}{t_0} - 1\right) \cdot \left(1 - \frac{t_{n-1}}{t_n}\right) \\ &= \frac{t_{n-1}}{t_n} \cdot \frac{t_n}{t_0} + 1 - \frac{t_{n-1}}{t_n} = 1 + \frac{t_{n-1}}{t_0} - \frac{t_{n-1}}{t_n} \end{aligned}$$

which concludes the proof. □

References

[1] M. Battaglini, Joint production in teams, J. Econ. Theory 130 (1) (September 2006) 138–167.
[2] Y.K. Che, S.W. Yoo, Optimal incentives in teams, Amer. Econ. Rev. 91 (2001) 525–541.
[3] Y. Eemek, M. Feldman. Computing optimal contracts in series-parallel heterogeneous combinatorial agencies, in: Proceedings of the Workshop on Internet and Network Economics, 2009, pp. 268–279.
[4] B. Holmstrom, Moral hazard in teams, Bell J. Econ. 13 (1982) 324–340.
[5] M. Feldman M. Babaioff, N. Nisan, Combinatorial agency, in: Proceedings of ACM Conference on Electronic Commerce, 2006, pp. 18–28.
[6] A. Ronen, L. Wahrmann, Prediction games, in: 1st International Workshop on Internet and Network Economics (WINE), 2005, pp. 129–140.
[7] I. Segal, Contracting with externalities, Quart. J. Econ. 114 (1999) 337–388.
[8] I. Segal, Coordination and discrimination with externalities: Divide and conquer?, J. Econ. Theory 113 (2003) 147–181.
[9] R. Smorodinsky, M. Tennenholtz, Overcoming free-riding in multi-party computations—The anonymous case, Games Econ. Behav. 55 (2006) 385–406.
[10] R. Smorodinsky, M. Tennenholtz, Sequential information elicitation in multi-agent systems, in: 20th conference on Uncertainty in Artificial Intelligence (AUAI), 2004, pp. 528–535.
[11] R. Strausz, Efficiency in sequential partnerships, J. Econ. Theory 85 (1999) 140–156.
[12] E. Winter, Transparency incentives among peers, RAND J. Econ. 41 (2010) 504–523.
[13] E. Winter, Incentives discrimination, Amer. Econ. Rev. 94 (2004) 764–773.