Transparency among Peers and Incentives*

Eyal Winter[†]

September 4, 2006

Abstract

The paper studies the effect of transparency among peers on the principal's cost of providing incentives. Using directed graphs to represent peer information we show that under complementarity the cost of providing incentives is decreasing with the level of transparency within the organization. We also investigate the role of the architecture of the information in boosting incentives. In arguing that substitution impedes the benefits of transparency we will compare function-based teams with process-based teams showing that the latter are more effective in providing incentives.

1 Introduction

The design of optimal incentives in organizations is closely tied to a variety of structural features of the workplace. A crucial one among those is the information structure among peers. A noticeable recent trend in the evolution of the workplace is the move away from private offices or cubicles to open-space environments or "war rooms." This policy which is implemented by different types of organizations is often referred to as "co-location." Coupled with other measures to increase interaction among peers, co-location enhances workers' monitoring opportunities and makes the information about effort

*I'm grateful to Ed Lazear, Bart Lipman, Phil Reny, and Ilya Segal for helpful suggestions. Comments and suggestions made by seminar participants at Boston U, Cornell, Ohio State, Bonn, Nottingham, Sydney, York, and Zurich are also gratefully acknowledged.

[†]The Center for the Study of Rationality and the Economics Department, The Hebrew University of Jerusalem, mseyal@pluto.huji.ac.il

more transparent among peers. The purpose of this paper is to study the effect of the internal information among peers on the design of optimal incentives in organizations. A variety of aspects influence peer information in organizations. One is indeed the workplace architecture, i.e., the extent to which workers operate in the line-of-sight of each other. But the structure of authority and the organizational culture may be equally relevant. We will use the term co-location in this paper to refer to any measure that makes peer information richer. A central question in our analysis would be whether, and under which conditions, more information among peers makes it easier for the principal to provide incentives.

Growing empirical evidence suggests that utilizing peer effects optimally is crucial to corporate success. Teasley et al. (2002) study the effect of co-location in software development teams. The authors evaluated the workers' productivity using measures commonly used in software development. Comparing the war room teams' scores with those in traditionally arranged offices, they conclude that teams in war rooms are twice as productive as similar teams working in closed offices. A related evidence is reported by Heywood and Jirjahn (2004) who show that blue-collar workers who work jointly in small teams have a lower absentee rate than other similar workers who work alone. An even cleaner evidence on the pure effect of observability on workers' incentives has been recently documented by Mas and Moretti (2006). Measuring productivity of checkers in a large grocery chain, they find that introducing a highly productive worker into a shift boosts the productivity of incumbent peers who are directly in the line-of-sight of the new worker. We will come back to these papers when pointing out some of the consistencies between our theoretical results and the empirical evidence.

To address the relationship between peer information and incentives we will present a moral hazard model in which agents' effort decisions are mapped into a probability that their joint project will succeed. The principal offers agents rewards that are contingent only on the final outcome of the project. Agents' effort while unobservable by the principal may be observable by peers. The internal information about peers will be given in a rather general form. Specifically, we will represent the information structure by a directed graph where an arc from agent i to agent j means that i is informed about $j's$ effort. We will interpret these graphs as emerging from the workplace structure where a movement towards co-location corresponds to richer graphs, i.e., more transparency among peers. Our analysis will compare different information structures in terms of the cost of the optimal incentive mechanism. In much of the paper we will distinguish between the case in which agents' tasks satisfy complementarity and the one in which they are substitutes. This distinction will allow us to address the issue of the optimal team formation in the second part of the paper.

Our first result asserts that under complementarity transparency among peers works in favor of the principal as it is easier to generate incentives under more transparent structures. Roughly, with more transparency among peers the implicit threat against shirking is stronger. Agents who are observed by many of their peers will be reluctant to shirk for fear that doing so will trigger the shirking of other peers, which will substantially reduce the probability that the project will succeed.

Not every information arc between two agents contributes to incentives. Proposition 3 asserts that the principal's cost of inducing all agents to exert effort depends on the graph (of information among peers) only through its transitive closure. Put differently, it does not matter whether agent i observes the effort of agent j directly or through observing other agents, i.e., observing an agent who observes an agent ... who observes agent j. In both cases the cost of inducing agent i to exert effort will be the same. This result will also allow us to compare workplace architectures on a wider domain than that of Proposition 1. The result implies that workplace structures that are rich in sequentiality are desirable from the point of view of incentives.

Architectural and organizational changes in the workplace affect the internal information peers have about each other. But in most environments this effect is not deterministic but rather probabilistic. A move towards co-location merely increases the probability that an agent will observe the effort of his peers. Section 6 is meant to address this issue with a model of random graphs. In this framework agent i observing agent j's effort is a random event that occurs with an exogenous probability. Our result in this section will show that the principal's cost of inducing effort by all agents is a decreasing function of this probability. Hence, even measures that only slightly increase the probability of transparency among peers will have a strict positive effect on incentives.

While peer information generates implicit incentives under complementarity Proposition 2 argues that when the technology is one of substitution co-location has no effect on incentives. The intuition behind the difference between complementarity and substitution relies on off-equilibrium behavior. Under complementarity agents are incentivized to shirk when observing the shirking of their peers which is crucial to generating the implicit threat against shirking. In contrast this threat disappears under substitution because under substitution a shirking agent generates greater incentives for his peers to exert effort. This intuition is consistent with a recent empirical study. Using panel data on the performance of baseball players Gould and Winter (2006) show that the nature of externalities among players depends on the degree of complementarity and substitutability between players. Complementary players induce positive externalities on each other; i.e., the effort of a player increases with the degree of effort by his peer. On the other hand players who are substitutes generate negative peer effects.

The distinction we will make between complementarity and substitution will allow us to address a related important issue: Should team formation be process based or function-based?[1] In a process-based team or war room each member is in charge of a different stage in the production process of a single product. In contrast, a function-based team accommodates agents who all work on the same stage of the production process (i.e., performing identical tasks on different products). We will argue that incentive considerations should favor process-based teams since agents' effort in such teams involves complementarity.

In the rest of the paper we discuss some extensions and generalizations. We consider an extended domain of information structures encompassing all imperfect information (extensive form) games. Our first result here shows that if we attempt to implement effort using Nash Equilibrium (rather than Perfect Bayesian Equilibrium), then the monotonicity result asserting "the more information among peers the better" holds for every technology (not just for those with complementarity) and it holds on the entire domain of information structures. We then identify an important condition that allows us to generalize Proposition 1 on the extended domain of information structure. This condition requires that substitution be ruled out in both the technology and the information structure, contributing to our basic intuition that substitution hampers the benefits of peer information.

This paper is part of the extensive literature on multi-agent incentive mechanisms that started with Holmstrom (1982). Some of the papers in this literature, such as Holmstrom and Milgrom (1990) and Itoh (1991), have pointed out that a principal can gain from collusion or coordination among

[1] I would like to thank Ilya Segal for pointing out the relevance of my model to this issue.

his agents. Baliga and Sjoestrom (1998) analyze a two-agent model allowing for a richer set of contracts. Agents can sign side contracts and the principal can contingent payments on messages. Their main focus is on the issue of delegation. The modeling choice in the current paper rules out both messages and side contracts on two counts. The first one is simplicity. The complexity of our model lies in the information structures and we try to keep contracts as simple as possible. The other reason is that, as stated earlier, our objective is to study the pure effect of information. Monitoring and being able to sign side contracts may end up being two different things and our choice in this paper is to focus on the first.

Closer to the spirit of our results is Che and Yoo (2001) who address a broader issue of comparing mechanisms based on relative performance to ones based on joint performance. They show that when two agents are involved in a repeated interaction and when the joint project has synergy a joint performance mechanism is superior as it generates an endogenous source of incentives through standard trigger strategies in a repeated game. Our results add to the important observation made by Che and Yoo (2001) in suggesting that the implicit incentives generated by team production may be effective not only under long-term interactions but also for one-shot activities.

Our point of departure from the literature discussed above is to a large extent defined by the objective of the paper, which is to study the pure effect of the architecture of information among peers on the cost of providing incentives. An important message our results add to this literature is that the mere fact that agents are informed about their peers' effort may be sufficient to boost incentives even when the organizational environment does not allow for real communication, the horizon is short, and contracts depend only on the project's outcome.

The moral hazard model we use here is similar to the ones used in two of my former papers. Winter (2004) argues that even when agents are identical and act simultaneously (i.e., with no information among peers) the principal may gain by discriminating among them. Winter (2005) considers sequential production and addresses the issue of how to optimally allocate tasks and agents with different attributes across different production slots. Recently some authors in computer science working on electronic markets have built on these papers to address the issue of combinatorial agency (see Babaioff, Feldman and Nisan (2006a) and (2006b)). Like almost all the related literature, these papers deal with a fixed structure of information about peers. In contrast, our main objective here is to compare between different information structures and study their effect on agents' incentives. Finally, we note that this paper, in certain respects, also relates to the recent literature on networks in games, which is extensively surveyed by Jackson (2005). Links in our networks stand for information channels between peers. However, in contrast to the networks literature, in which the game involves forming the network, here the network is designed by the principal and the interaction among the agents relies on the network rather than forms it.

2 A Simple Two-Agent Example

Before introducing the formal model in its generality it would be instructive to demonstrate the incentive gains from transparency among peers with the simplest possible example. Consider an organization of two agents, each of whom is in charge of one task on which he can either exert effort or shirk. The cost of effort is $c > 0$ for both agents. If an agent does not exert effort in his task the task will end successfully with probability $\alpha > 0$. If the agent exerts effort the task will succeed with probability 1. The two tasks together form a project that will succeed if and only if both tasks are successful. The principal cannot monitor the agents' effort or the outcome of the individual tasks. He can only find out whether the entire project is successful. A mechanism pays the agents rewards v_1, v_2 if the projects succeeds and zero if it fails. We would like to find the mechanism with a minimal total reward that will induce both agents to exert effort in equilibrium.

We first assume that the two agents act simultaneously; i.e., neither is informed of the effort decision of the other. For effort by both players to be an equilibrium it must be that effort by i is his best response to effort by j. Suppose that agent i exerts effort, then agent j's payoff will be $v_j - c$ if he exerts effort as well and it will be αv_j if he shirks. We therefore must have $v_j - c \geq \alpha v_j$. Hence the principal must pay $c/(1-\alpha)$ to each agent in order to sustain effort by both.

Consider now the case in which the agents act sequentially; i.e., agent 1 makes his effort decision. Agent 2 observes that decision (but not the outcome of the task) and decides whether to exert effort or to shirk. To generate an equilibrium in which both agents exert effort player 2 must be paid $c/(1-\alpha)$ (as in the simultaneous case). If he is paid less his best response to seeing agent 1 exerting effort is to shirk. How much then should agent 1 be paid to sustain an effort equilibrium? Suppose that agent 1 is promised

$v_1 = c/(1-\alpha^2)$ if the project succeeds and player 2 is promised $c/(1-\alpha)$. We will argue that there exists an equilibrium in which both agents exert effort under these rewards. Consider the strategy profile under which agent 1 exerts effort and agent 2 exerts effort if and only if agent 1 does so. If agent 1 deviates from his strategy by shirking agent 2 will shirk as well and agent 1 will get $\alpha^2 v_1$. If agent 1 instead exerts effort the project will succeed with probability 1 and he will get $v_1 - c$. Since $\alpha^2 v_1 = v_1 - c$ it is optimal for agent 1 to exert effort, and given that agent 1 exerts effort it is optimal for agent 2 to do so. In fact, for an arbitrarily small increase of the rewards for the two agents the strategy profile specified above is also the unique subgame perfect equilibrium of the game as player 2 will indeed find it optimal to shirk if he observes player 1 doing so.

Because $v_1 = c/(1-\alpha^2) < c/(1-\alpha)$ the principal needs to pay less if the agents move sequentially. Put differently, the fact that agent 2 is informed of agent 1's effort decision makes it easier for the principal to provide incentives for effort. Roughly, the implicit threat that agent 1 faces not to shirk (because agent 2 will do so), which allows the principal to reduce rewards for agent 1, applies only when they move sequentially[2]. We will build on this example in constructing a general model of an organization with internal information about effort.

3 The Model

The organizational project involves a set N of n agents who collectively manage a project. Each agent has to decide whether to exert effort/invest in the performance of his tasks or not. The cost of effort is c and is constant across all players. Henceforth we interchangeably use the term *investment* to mean the action of exerting effort. The technology of the organization maps a profile of effort decisions into a probability of the project's success. For a group $S \subseteq N$ of investing agents the probability that the project will succeed is $p(S)$. Throughout the paper we assume that p is increasing in the

[2] As mentioned earlier we are ruling out messages and side contracting as used in Baliga and Sjostrom (1998). It is however interesting to point out that if agents can send messages to the principal then the optimal contract in the sequential game would be $v_1 = c$ and $v_2 = \frac{c}{1-\alpha}$ (with agent 2 blowing the whistle when agent 1 shirks). This makes the gap between sequentiality and simultaneity even wider. This contract also seems to be collusion proof because joint effort dominates joint shirking as $v_1 + v_2 \geq \frac{2c}{1-\alpha^2}$.

following simple sense: if $T \subset S$, then $p(T) < p(S)$.

The principal who cannot monitor the agents for their effort but knows only if the project succeeds sets up a mechanism $v = (v_1, ..., v_n)$ by which agent i receives the payoff v_i if the project succeeds and zero otherwise.

Agents' decisions about whether or not to invest depend on the internal information about other agents' effort decisions.

An *Internal Information about Effort* (IIE) is defined to be a partial order "k" over the set of agents N, where $i\ k\ j$ stands for agent i knowing the effort decision of agent j before making his own decision. We will also refer to this relation by saying that "i sees j".

We impose that k is acyclic; i.e., for any sequence $i_1\ k\ i_2\ k, ..., k\ i_r$ with $r \leqslant n$ we must have that $i_1, ..., i_r$ are distinct. This condition simply reflects the fact that there can be no mutual knowledge about effort. Indeed $i\ k\ j$ implies that j has taken his effort decision before i, which precludes $j\ k\ m$ for some m with $m\ k\ i$. The acyclicity property of k also implies that there is a timing structure according to which agents make their investment decisions. We denote by K_i the set of agents that i sees, i.e., $K_i = \{j |\ i\ k\ j\}$.

Given a vector of rewards $v = (v_1, ..., v_n)$ that are to be paid if the project succeeds, and an IIE k, we can now consider the following game $G(k, v)$:

A strategy for player i is a function s_i: $2^{K_i} \longrightarrow \{0, 1\}$ specifying to each player whether to invest (choose 1) or to shirk as a function of the information he possesses on other agents' decisions. For every strategy profile $s = (s_1, ..., s_n)$ we denote by $M(s)$ the set of agents who exert effort under the profiles s.[3] Finally, the payoff for player i under $s = (s_1, ..., s_n)$ is given by $f_i(s) = v_i p(M(s)) - c$ if $i \in M(s)$ and $f_i(s)= v_i p(M(s))$ if $i \notin M(s)$.

We say that the vector of rewards v is an *investment-inducing (INI)* mech-

[3] We point out that the set $M(s)$ is defined inductively. Define by $k_i = \{j \in N | j\ k\ i\}$, i.e., the set of agents seeing i. We say that i is in the core of k if k_i is empty. Since k is acyclic it has a non-empty core which we denote by $C(N)$. Given the strategy profile s, actions are uniquely determined for all players in the core. Consider now the binary relation k restricted to the set of players $N \backslash C(N)$. This binary relation is again acyclic and has a non-empty core which we denote by $C(N \backslash C(N))$. Any player j in $C(N \backslash C(N))$ is informed only of actions taken by players in $C(N)$. Hence, given that the actions taken by players in $C(N)$ are well defined and unique, so are the actions of players in $C(N \backslash C(N))$. We can now proceed by induction. At each stage we eliminate players for which the actions have already been determined and we remain with an acyclic graph on the remaining players, which has a non-empty core. The process terminates when no players are left, at which stage we attain the vector of actions consistent with the profile s, and $M(s)$ is simply the set of players who choose to exert effort.

anism with respect to k if it induces all agents to invest in equilibrium. Formally, v is an INI mechanism if there exists a *Perfect Bayesian equilibrium (PBE)* s for the game $G(k, v)$ with $M(s) = N$.

We say that v is an *optimal* INI mechanism if it yields minimal total payoff among all INI mechanisms.

For an IIE k we denote by $v^*(k)$ the total reward in an optimal INI mechanism with respect to k. We point out that the assumption that the principal provides incentives to *all* agents is for the sake of simplicity. An alternative model in which the principal attempts to maximize the net profit of the project (taking into account the value of the project) makes the analysis more cumbersome and offers no additional insight.

As indicated earlier the binary relation "k" represents the prevailing information structure about peers' effort in the organization. Our objective is to compare such information structures in terms of the cost of inducing effort by all agents. To this end we will define a binary relation over IIEs by which we will be able to say that the one contains more information about effort than the other.

Let K be the set of all IIEs. For k_1 and k_2 in K we say that k_1 is richer than k_2 if for all i, j in N we have $i\ k_2\ j$ implies $i\ k_1\ j$. We shall also refer to k_1 as representing more transparency among peers than k_2.

4 Complementarity vs. Substitution

Throughout most of the paper we will distinguish between the properties of complementarity and substitution of the organizational technology. We say that a technology p satisfies complementarity across agents if for every two sets of agents (coalitions) S, T with $T \subset S$ and every agent $i \notin S$ we have $p(S \cup \{i\}) - p(S) > p(T \cup \{i\}) - p(T)$, i.e., i's effort is more effective as the set of other agents who exert effort increases. In contrast, we say that p has substitutability across agents if $p(S \cup \{i\}) - p(S) \leq p(T \cup \{i\}) - p(T)$. Note that the example in Section 2 is one of complementarity.[4] Specifically, $p(S) = \alpha^{2-|S|}$. Proposition 1 argues that if the technology p satisfies complementarity, then it is easier to provide incentives under more transparency within the organization.

[4] Complementarity (resp. substitution) can be associated with the convexity (resp. concavity) of the technology p.

Proposition 1: Let p satisfy complementarity. If k_1 is richer than k_2, then $v^*(k_1) \leqslant v^*(k_2)$.

We will omit the proof of Proposition 1 here as we will later state and prove a stronger result. However, the example in Section 2 suggests the intuition. In a richer IIE effort is more transparent, which allows more agents to make their effort decision dependent on that of a greater number of their peers. For each of these agents shirking will result in a more detrimental effect on the project's success probability and thus also on the expected reward this agent receives. Hence, the exposure of an agent's effort increases his/her incentive to exert effort and thus decreases the principal's cost of incentivizing that agent. The complementarity condition assures that the strategy profile that sustains the implicit incentives is sequentially rational in the sense that agents will indeed shirk if they observe others shirking.

The relevance of Proposition 1 and the above intuition to real-life organizations is nicely documented in an interesting case study by Knez and Simester (2001). In February 1995, on the verge of bankruptcy, Continental Airlines (CA) introduced the Go Forward Plan to improve on-time departure performance by offering \$65 to every CA employee in every month that CAs on-time performance is ranked in the top 5. It is claimed that this plan was central in moving the company from a \$613m loss for 1994 to a profit of \$224m for 1995. The authors argue that crucial to the success of this plan were the following two factors: 1. Much of the activity affecting on-time performance in CA was carried out in teams of small groups making the information about peers' effort transparent. 2. The technology of improving on-time performance is one of complementarity, which makes peer information effective.

In contrast to the case of complementarity under substitution across agents information among peers turns out to have no effect on incentives. If the principal wishes to implement full effort he will have to provide sufficient incentives for agents to exert effort when they believe that all other agents do so as well. But this means that because of substitution whenever an agent observes some of his predecessors shirking his only rational action is to exert effort. This sterilizes the implicit incentive that makes an agent cautious about shirking in the complementarity case thus making the information structure irrelevant for incentives. Formally we can show

Proposition 2: Let p satisfy substitution across agents. Then the optimal INI mechanism is identical for all information structures and pays agent i $v_i^* = c/[p(N) - p(N \backslash \{i\})]$.

Proof: Assume that p satisfies substitution across agents. For each agent i let $v_i^* = c/[p(N) - p(N\backslash\{i\})]$. We argue that v_i^* is the optimal incentive-inducing mechanism regardless of the information structure about peers' efforts. Indeed, consider the strategy combination in which each agent exerts effort regardless of the information that he/she observes about other agents' effort. We argue that if agents' rewards are given by v_i^* then this strategy profile is a Perfect Bayesian equilibrium regardless of the information structure under which the game is played. To see that it is a Nash equilibrium note that under v_i^* each agent is indifferent between exerting effort and shirking given that all the other players are investing (i.e., v_i^* solves $vp(N)-c = vp(N\backslash\{i\})$). To show that the equilibrium is PBE we attribute the following set of beliefs to agents: At a given information set each agent believes that his predecessors whom he could not observe are exerting effort. Suppose that an agent observed the set of agents S shirking. Given his beliefs and given that the continuation of the game is carried out according to the above profile, agent i's expected payoff when he exerts effort is $v_i^* p((N\backslash S) \cup \{i\}) - c$ and is $v_i^* p(N\backslash S)$ when he shirks. But since p satisfies substitution across tasks we have $p(N) - p(N\backslash\{i\}) \leq p((N\backslash S) \cup \{i\}) - p(N\backslash S)$ and i is better off exerting effort. Finally, we note that the mechanism v_i^* is optimal. Consider some agent i and assume a different mechanism that pays $v_i' < v_i^*$ to player i. Suppose by way of contradiction that an equilibrium exists in the corresponding game with all players exerting effort. It must be the case that this equilibrium specifies that each player exerts effort if he encounters no shirking. Otherwise, along the equilibrium path there will be at least one agent shirking. But because all agents are indifferent between investing and shirking under v_i^*, they will be induced to shirk under v_i'. Q.E.D.

Proposition 2 indicates that substitution impedes the benefits of peer information. One may wonder, however, why the principal cannot be strictly worse off with more information among peers. In Appendix 1 we show by means of an example that this can indeed be the case if there are multiple levels of effort and when the principal attempts to incentivize agents to exert less than full effort. We then provide a two-agent result showing that in such a framework the principal is better off with more information under complementarity and is (weakly) worse off under substitution.

As we pointed out the complementarity condition in Proposition 1 guarantees that player who observes their peers shirking will find it optimal to shirk as well. Hence the trigger strategies that sustain effort are optimal both on and off the equilibrium path. This suggests that if we intend to implement effort using Nash equilibria (rather than PBE) Proposition 1 will hold for every technology, which is precisely the statement of Proposition 3. We defer the proof until later when we state still a stronger result.

Proposition 3: Suppose that the optimal mechanism is defined based on Nash equilibrium (rather than PBE). Then for any technology p, if k_1 is richer than k_2, then $v^*(k_1) \leqslant v^*(k_2)$.

From a theoretical point of view one may argue that PBE is a more appropriate solution concept than Nash equilibrium for implementation in our context because of the sequential structure of the game. However, from an empirical point of view Nash equilibrium implementation seems to be of considerable importance too. There is overwhelming empirical evidence on peer effects suggesting that trigger strategies of the type we described above are prevalent in various organizations including ones which involve tasks that are not clearly complementary (see for example Ichino and Maggi (2000), which discusses the way workers in large Italian banks affect each other effort, as well as Mas and Moretti (2006) which we discussed earlier.) One can indeed attribute these peer effects to behavioral motives such as reciprocity and fairness (see Kandel and Lazear (1992)) but it is important to note that they are also consistent with Nash equilibrium strategies without any behavioral assumptions.[5]

At this point and unless otherwise specified we will confine ourselves to technologies satisfying complementarity.

5 Indirect Information about Effort

As we have seen in Proposition 1 the network of information about peer effort affects the principal's cost of providing incentives. We will now to go a step further in understanding how the information structure affects incentives by extending Proposition 1 to the case where IIEs are not necessarily comparable in terms of "richness." This analysis relies on the observation that agent i's shirking may also affect the behavior of agents who do not observe i's shirking directly. Suppose for example that agent 1's action is observed only by agent 2, while agent 2's action is observed only by agent 3. Assuming that the three agents adopt strategies by which they shirk if and only if they observe

[5] In a similar context several papers on the ultimatum bargaining make the point that the concept of Nash equilibrium is empirically more relevant than subgame perfection in spite of the latter being theoretically more appealing (see for example Roth et al. (1991)).

one of their peers shirking, we will see agent 1 affecting agent 3's behavior without agent 3 being directly informed of what agent 1 did. In designing the optimal mechanism for a given IIE the principal will attempt to sustain as an equilibrium the behavior described above by which agents exert effort unless they detect one of their peers shirking. This in particular means that enriching the information structure described in the above example by letting agent 3 directly observe agent 1's effort choice does not allow the principal to reduce the cost of incentivizing agent 1. This all hints at the fact that in the IIE only those arcs $i\ k\ j$ matter for which there is no way for i to learn indirectly about the action of j through the actions of other peers. This leads us to the graph-theoretic notion of Transitive Closure.

For an IIE k we denote by $t(k)$ the IIE obtained from the transitive closure of k. Specifically, in the IIE $t(k)$ we have $i\ t(k)\ j$ if and only if there exists a sequence of agents $i_1, i_2, ..., i_r$ with $i_1 = i$ and $i_r = j$ and $i_m\ k\ i_{m+1}$ $(m = 1, 2, ..., r-1)$. Proposition 4 asserts that from the point of view of incentives only the transitive closure of an IIE matters.

Proposition 4: Let p be a technology satisfying complementarity. If $t(k_1)$ is richer than $t(k_2)$, then $v^*(k_1) \leqslant v^*(k_2)$, with strict inequality if and only if $t(k_1) \neq t(k_2)$.

Proof: For each $i \in N$ let $C(i, k)$ be the set of agents seeing i according to $t(k)$, i.e., $C(i, k) = \{j \in N|\ j\ t(k)\ i\}$. We will show that $v_i = \frac{c}{p(N)-p(N\backslash[C(i,k)\cup\{i\}])}$ is the optimal incentive-inducing mechanism. Note that the payoff v_i makes i indifferent between shirking and investing provided that his shirking triggers the shirking of all j in $C(i, k)$ and no more; i.e., v_i solves the following:

$p(N)v_i - c = v_i p(N\backslash[C(i, k) \cup \{i\}])$. (*)

If $t(k_1)$ is richer than $t(k_2)$, then $C(i, k_2)$ is a subset of $C(i, k_1)$ for all i. Furthermore, if $t(k_1) \neq t(k_2)$, then for some i this inclusion is strict. Hence the mechanism proposed above satisfies $v^*(k_1) < v^*(k_2)$. It thus remains to show that the mechanism described above is the optimal incentive-inducing mechanism. We first argue that with the above mechanism a Nash equilibrium exists in which all agents invest. We then show that this equilibrium is also a (PBE). Consider the strategy profile in which each agent invests unless he sees (directly with respect to k) someone shirking. Under this profile all agents invest. Furthermore, any deviation by player i will trigger (sequentially) all players in $C(i, k)$ to shirk. Hence, if i shirks his expected payoff will be according to the RHS of equation (*). This means that i cannot increase his payoff by deviating. So the profile above forms a Nash equilibrium. The strategy profile described above is also a (PBE) supported by the following set of beliefs: Each agent i facing a certain information set believes that his predecessors whose actions he could not observe are exerting effort. In particular if he observes no shirking he believes that everyone exerted effort. To show that this set of beliefs together with the strategy profile described earlier forms a (PBE) we need to show that given the players' beliefs and provided that the continuation game is carried out according to the profile described above each agent is acting rationally at every possible information set. Clearly, for the information set in which player i observes no one shirking and given the specified beliefs this boils down to the inequality $p(N)v_i - c \geq v_i p(N \backslash [C(i,k) \cup \{i\}])$, which holds because of (*). Hence, it is enough to show that player i acts rationally at information sets in which he observed someone shirking. Let R be the set of agents that i observes shirking at such an information set. Denote by $R^* = \{j \in N | j\ t(k)\ m\ for\ some\ m \in R\}$; i.e., R^* is the set of agents who see someone that sees someone ... who is in R. Hence, based on i's beliefs, the shirking of the agents in R will trigger all the agents in R^* to shirk even if i himself is exerting effort. The expected payoff of agent i in this case will be $v_i p(N \backslash R^*) - c$. On the other hand, if i shirks his payoff is $v_i p((N \backslash R^*) \backslash (C(i,k) \cup \{i\}))$ (i.e., in addition to R^* the coalition $C(i,k) \cup \{i\}$ will also shirk). At this stage we need a bit of set theory algebra:

Lemma: If p satisfies complementarity across tasks, then for any two non-empty coalitions of agents $R, C \subset N$ we have $p(N) - p(N \backslash C) > p(N \backslash R) - p((N \backslash R) \backslash C)$.

Proof: We first note that if p satisfies complementarity, then the inequalities defining it also apply to coalitions; i.e., for any two coalitions S, T with $T \subset S$ and a coalition Q with $Q \cap S = \emptyset$ we have

$p(S \cup Q) - p(S) > p(T \cup Q) - p(T)$ (**).

This is proved by induction on the number of members in Q using the definition of complementarity agent by agent. Let $T = (N \backslash R) \backslash C$, $S = N \backslash C$ and $Q = C$. Because $C \neq R$ we have $T \subset S$ and we can use (**) to establish the claim of the lemma.

By the lemma above we conclude that $p(N) - p(N \backslash [C(i,k) \cup \{i\}]) > p(N \backslash R^*) - p((N \backslash R^*) \backslash (C(i,k) \cup \{i\}))$ (i.e., setting $R = R^*$ and $C = C(i,k) \cup \{i\}$) and thus $v_i p((N \backslash R^*) \backslash (C(i,k) \cup \{i\})) > v_i p(N \backslash R^*) - c$. Hence, player i is better off shirking at such an information set showing that the specified equilibrium is (PBE).

It remains to show that any mechanism that pays less than v_i to some player cannot admit an equilibrium in which all agents invest. Indeed, consider by way of contradiction an equilibrium of the underlying game in which all agents invest. Consider such a player i. If $C(i,k)$ is empty, then i is affecting no player's behavior. Since i is indifferent between investing and shirking in v_i, a lesser payoff will make shirking strictly preferable. Assume now that $C(i,k)$ is not empty, then shirking by i will trigger at most the set of players $C(i,k)$ (and possibly less) to shirk as well. Since v_i makes i indifferent between investing and shirking, a lesser payoff will make shirking strictly better, which contradicts the equilibrium assumption. Q.E.D.

Proposition 4 suggests that architectures with a high degree of sequentiality are more conducive to incentives. In particular it implies that simply counting arcs may be misleading. For example consider graphs A and B in Figure 1. Graph A has more arcs than graph B. However, the transitive closure of A is the graph A itself, while the transitive closure of B is the graph C which is richer. Hence the cost of providing incentives under graph B is lower than under graph A.

An almost immediate consequence of Proposition 4 is that an information structure corresponding to a chain is optimal for the principal. Specifically, an IIE k is said to be a chain if there exists an order of the agents $i_1, i_2, ..., i_n$ such that the set of arcs in k is given by $i_1 k\ i_2, ..., k\ i_n$. Such an IIE can be interpreted as a sequential production in which each agent observes the effort decision of his immediate predecessor. Furthermore, it follows trivially from Proposition 4 that the empty IIE, which corresponds to no internal information about effort at all (as if all agents decide their effort simultaneously), is the least attractive for the principal.

Corollary 1: If $v^*(k_s) \leqslant v^*(k)$ for all k, then the IIE k_s must correspond to a chain.

Proof: The result follows from the following graph-theoretic claim: If the graph g is a chain, then the transitive closure of g is a maximal graph (richest) among all acyclic graphs. This can be shown as follows: Take an ordered sequence of nodes $1, 2, ..., n$, and consider the graph $g^* = \bigcup_{i=1}^{n-1} \{(i,k)|\ k > i\}$, where (i,j) denotes an arc from i to j. It is easy to note that g is maximal among all acyclic graphs. Indeed, if we add another arc to g it must be of the form (i,j), where $j < i$ and $2 \leq i \leq n$. But upon addition of such an arc we create the cycle $(i,j),(j,i)$. Hence g is maximal among all acyclic graphs. Consider now the chain graph given by $g = \{(1,2),(2,3),...,(n-1,n)\}$. It is easy to see that the transitive closure of g is $g^*$and the claim follows. The corollary now follows directly from the claim together with Proposition 4.

Q.E.D.

Corollary 2: Among all IIEs the IIE k_e that corresponds to the empty graph yields the maximal cost for the principal, i.e., $v^*(k_e) > v^*(k)$ for all k.

6 Random Peer-Monitoring

The fact that the internal information about effort among peers affects the principal's cost of incentivizing his agents suggests that the principal might want to influence this information in order to reduce his cost. Since Proposition 1 asserts that the principal can only gain from more transparency among his peers he might want to "enrich" the IIE so that more agents would be informed about more of their peers. This can be done, for example, by forming teams and locating workers in one office. It can also be achieved by organizing more frequent recreational activities within the organization through which the information about effort is transmitted more effectively. Finally, it can even be promoted by simply redesigning the physical architecture of the workplace so that there is more visibility across workers. Of course, when promoting the objective of more visibility among peers the principal will be sensitive to other managerial constraints that are not incentive-based (such as the cost of reorganization). But regardless of the method the principal chooses to adopt in order to facilitate more transparency among peers it is reasonable to assume that whether agent i is informed about agent j's effort remains a random event. The tools and means by which the principal can affect the transparency among his peers can only influence the probability that this event will take place. We would therefore like to move on now to a model in which agents' monitoring opportunities are random. Specifically, we would like to model a situation where agents are uncertain about which of their peers can observe their effort decisions. Roughly speaking, we will argue that the principal's cost of incentivizing his agents is monotonically decreasing with the probability with which an agent can observe his peers. Thus any measure taken by the principal that can increase even slightly the probability of the availability of information will result in a reduction in the principal's cost of providing incentives.

To be able to address this issue we will need to separate between the chronology of the order of moves and the other factors that determine the information about peers. This will be done by making the order of moves explicit.

A *random* IIE is defined by a pair (w, k_q), where w is an order of the agents and k_q is a random directed graph on the set of agents with arcs emerging randomly according to an IID Bernoulli distribution with probability $0 < q < 1$ (i.e., every arc forms with probability q independently of the other arcs). We interpret w as the order in which agents take their effort decisions, and k_q represents the random structure of the information about peer effort. An arc from i to j represents $i's$ technical or formal ability to monitor agent j's effort. More specifically, agent i is informed of agent's j's effort decision before making his own if and only if i appears after j in the order w and the arc (i, j) from i to j is realized in k_q. When agent i takes his turn to act (in the order w) a realization takes place to determine which of the arcs (i, j) with $w(j) < w(i)$ forms, i.e., whose effort action does i see among those who preceded him? However, at this stage i is uncertain about who is going to observe his own effort decision. We assume that each agent is informed about the arcs formed among his predecessors but as we will see this assumption is inessential to the analysis. Of course the pair (w, k_q) is commonly known to all agents. The principal himself has to design the incentive mechanism *ex ante* before any realization. He is informed only about q and w. As before, given a reward mechanism v and a random IIE (w, k_q) agents are facing a game (now involving Nature's random moves). The definition of an optimal incentive-inducing mechanism remains unchanged.

For a random IIE (w, k_q) we denote by $v^*(w, k_q)$ the total reward of the optimal incentive-inducing mechanism under k_q. In the sequel and with a slight abuse of notation we will use $v^*(q)$ for $v^*(w, k_q)$. The result by which the principal should favor transparency among peers now reads as follows:

Proposition 5: If $q' > q$, then $v^*(q) > v^*(q')$.

The proof is relegated to Appendix 2.

7 Function-Based vs. Process-Based Teams

The distinction between complementarity and substitution that we made in Propositions 1 and 2 not only identifies the environments in which peer information is expected to be more effective, but it also has an important implication for the optimal design of teams, i.e., which agents should be located together to share information? Being aware of the fact that other managerial considerations other than incentives may favor keeping teams small, we would like to raise the issue of how to design teams optimally given a size constraint. We shall consider an environment in which there are several products each of which is produced through a process in which different agents assume different functions. In a function-based team all agents perform the same function on different products. A process-based team involves agents dealing with the entire process of a single product, each agent assuming a different function. Suppose for example that a project involves the preparation of nine turkey sandwiches each of which has to undergo a process of cutting the bread, spreading the mayonnaise, and slicing the turkey. There are nine workers of which three are bread cutters, three are mayonnaise spreaders, and three are turkey slicers. A process-based structure will have three teams each consisting of one bread cutter, one mayonnaise spreader, and one turkey slicer. A function-based structure will again have three teams one consisting of all the bread cutters, one of only mayonnaise spreaders, and one of only turkey slicers.

Our objective in this section is to argue that incentive considerations should favor the process-based structure to the function-based one.

The intuition is roughly as follows: Agents who assume different functions within the production of the same product are involved in tasks that are complementary, whereas the relation between agents performing the same function is that of substitution. As we argued earlier, agents with complementarity can generate implicit incentives when possessing information about peer effort, while agents with substitution cannot. This means that in utilizing the implicit incentives optimally the principal should co-locate agents among which there is complementarity.

We will now provide a formal argument to our intuition above, which, for simplicity, will be presented in the framework of two-function, two-product organizations. We first need to set up the properties of complementarity and substitution as a binary relation among agents.

We say that agents i and j are complementary with respect to the technology p if the investment of each of these agents is more effective when the other agent invests than when he does not. Formally: For each coalition of agents S such that $i, j \notin S$ we have $p(S \cup \{i, j\}) - p(S \cup \{j\}) > p(S \cup \{i\}) - p(S)$. We say that agents i and j are substitutable if each agent's investment (weakly) decreases the marginal contribution of the other agent, i.e., $p(S \cup \{i, j\}) - p(S \cup \{j\}) \leq p(S \cup \{i\}) - p(S)$.[6]

[6] See Segal (2003) who discusses similar properties to study integration among agents

Consider now the following four-agent organization involving the production of two products A and B. The production of each product involves a two-stage process: an upstream stage and a downstream stage. Each agent is in charge of one production stage of one of the products. We denote by a_d, a_u, b_d, b_u the four agents according to the tasks they are in charge of (i.e., a_d is in charge of the downstream task of product A, etc.). We assume that there exists complementarity across different stages of the same product and substitution across different products at the same stage. Specifically, there exists complementarity between a_d and a_u as well as between b_d and b_u, whereas a_d and b_d are substitutes and so are a_u and b_u. An example of a technology satisfying these conditions can be given as follows: Assume that each stage of production succeeds with probability α if effort is not exerted and with probability β with $\beta > \alpha$ if effort is exerted. Each product succeeds if and only if the two stages of its production end successfully. Define now the project's goal to be the successful production of at least one of the two goods. The resulting technology p satisfies precisely the conditions imposed above.

Suppose now that teams are constrained to contain no more than two agents. The process-based structure involves two teams $\{a_d, a_u\}$ and $\{b_d, b_u\}$, while the teams in the function-based structure are $\{a_d, b_d\}$ and $\{a_u, b_u\}$. To introduce the effect of co-location on agents' information about peers' effort we assume that in both structures and in each of the teams one agent observes the action of the other agent in his/her team before performing his/her own (i.e., an a-type acts before a b-type, and a d-type before a u-type). No effort information is revealed between the teams. We can now assert the following:

Proposition 6: The optimal mechanism in the process-based structure costs less than the optimal mechanism in the function-based structure.

The proof is relegated to Appendix 2.

8 General Information Structures

In this section we will consider general information structures among peers in which agents can be exposed to more complex signals on peers' effort. The purpose of this section is twofold. Firstly, it will demonstrate that our results about the effect of peer information extends to more general information structures. Secondly, it will allow us to test the boundaries of these results.

in a cooperative game using a Shapley value framework.

We maintain our separate treatment of the special case of directed graphs as it allows us to deal with network architectures and indirect monitoring in a way that we cannot do in the general case.

Our first result here will show that under Nash implementation our monotonicity result (Proposition 1) becomes extremely general. It holds for any technology and for all information structures within our general domain. But our treatment of general information structures sheds an interesting light on the limit of our monotonicity result under Perfect Bayesian implementation. It turns out that for the same reason we needed lack of substitution in the technology to establish our monotonicity result (see Proposition 1) we also need lack of substitution in the information structures for the result to hold. We will first provide an example showing that the monotonicity result cannot be extended to the entire class of information structures, and will immediately then restore the result by imposing a condition which we call "lack of informational substitution." This condition roughly requires that an agent be unable to conceal the shirking of his/her peer from a third agent by substituting for his/her effort. We start by describing the general model.

Agents move sequentially in deciding on their effort (player 1 moves first and player n moves last).

For a given vector of rewards the players play an extensive form game which is given by a binary tree. At each decision node a player has a binary choice of 1 (for effort) or 0 (for shirking). Player m has 2^{m-1} decision nodes in the tree, each of which is a binary vector of size $m-1$ specifying the actions taken by m's predecessors. We denote such decision node by x_m and use D_m to denote the set player m's decision nodes. We denote by $\mathbf{1}= (1, 1, ..., 1)$ the decision node of m that corresponds to all predecessors exerting effort (i.e., choosing 1). An information structure I is now given by a sequence of partitions $I = (P_1, ..., P_n)$, where P_m is a partition of D_m. We will say that the information structure I^1 is more transparent than I^2 if for all $1 \leq m \leq n$ we have P^1_m is a more refined partition (in the weak sense) than P^2_m. (Formally, for each set of nodes $S \in P^1_m$ there exists a set of nodes T with $S \subset T$ and $T \in P^2_m$.) We denote by I^* the most transparent information structure, i.e., the one in which each information set consists of a single decision node (perfect information).

The notion of optimal mechanism is defined now in precisely the same manner as before. For each information structure I and a technology p we denote by $v^*(I, p)$ the cost of the optimal mechanism under the information structure I and the technology p. We can state the following proposition which we prove in Appendix 2:

Proposition 7: Let I^1 and I^2 be two information structures such that I^1 is more transparent than I^2. Then under Nash implementation for *any* technology p we have $v^*(I^1,p) \leq v^*(I^2,p)$.

To establish a result in the spirit of Proposition 7 also for the case of Perfect Bayesian implementation we would need first to require that the technology is one of complementarity (see Proposition 1). However without an additional requirement on the information structures the monotonicity may fail to hold. Interestingly, in addition to ruling out technological substitution one would have to rule out informational substitution to establish the desired monotonicity result. Roughly speaking, the condition requires that a player cannot hid the shirking of other players by substituting for their effort. We first demonstrate that absent such a condition more transparency may result in a more expensive mechanism (i.e., a higher cost for the principal).

Example

Let I^1 be the information structure in which player 2 sees the effort decision of player 1 and players $3, 4, ..., n$ are informed only when both player 1 and player 2 are shirking (and know nothing about each other). Let I^2 be the information structure in which player 2 is again informed about the effort decision of player 1 but players $3, 4, ..., n$ receive no information whatsoever. Clearly I^1 is more transparent than I^2.

Claim: For a technology p satisfying complementarity it is easy to verify that $v^*(I^1,p)$ pays $v_j^* = \frac{c}{p(N)-p(N\backslash j)}$ for $j = 1, 2, ..., n$ while $v^*(I^2,p)$ pays $v_1^* = \frac{c}{p(N)-p(N\backslash\{1,2\})}$ and $v_j^* = \frac{c}{p(N)-p(N\backslash j)}$ for $j = 2, 3, ...n$. Hence, the more transparent information structure induces more cost to the principal. The intuition behind the claim and the violation of monotonicity in this example is quite clear. In the structure I^1 player 2 can hide (from the rest of the players) the fact that 1 shirked by exerting effort himself. Indeed, under such circumstances he would choose to do so otherwise he would trigger the shirking of all the remaining players. This means that player 1 loses the implicit incentive to exert effort as he knows that player 2 will substitute for him. Consequently the principal must boost the reward for player 1. In contrast, if $3, 4, ..., n$ know nothing about 1 and 2, then 2 would have no incentive to substitute for 1 and the implicit incentive for player 1 is in place (requiring a smaller reward by the principal).

We now impose a condition on information structures under which the monotonicity can be restored in Perfect Bayesian implementation.

Let x and x' be two decision nodes (paths) of player m. We denote by $y = x \wedge x'$ the decision node with $y(j) = \min\{x(j),\ x'(j)\}$. In words, player j (who moves before m) shirks on y if he shirks on at least one of the paths x and x'. We say that the information structure I *lacks informational substitution* (LIS) if $x, x' \in P_m(\mathbf{1})$ implies that $x \wedge x' \in P_m(\mathbf{1})$ for every player m and every pair of decision nodes x and x' in D_m. Note that x and x' are two play paths that specify the actions taken before m moves. The condition above says that if m cannot verify that someone shirked following the play of either paths x and x' then he also cannot verify that someone shirked following a path in which the set of shirking players is the union of those who shirk in x and x'. Clearly this condition is satisfied for information structures which are based on a directed graph, but the class of structures satisfying this condition is much larger. For example an information structure that provides the number of shirking agents without revealing their identity satisfies the (LIS) condition. Finally note that the information structure I^1 in the above example does not satisfy (LIS) as the condition fails for every player $m \geq 3$.

Proposition 8: Let I^1 and I^2 be two information structures such that I^1 is more transparent than I^2 and such that both satisfy lack of informational substitution. Then under Perfect Bayesian implementation for any technology p with increasing returns to scale we have $v^*(I^1, p) \leq v^*(I^2, p)$.

See Appendix 2 for the proof.

9 Discussion

We have seen that under complementarity and for a large domain of information structures, more transparency among peers makes it easier for the principal to provide incentives. Environments in which workers perform distinct tasks within a project whose output relies on the success of its weakest link seem to have this property. Software development, R&D projects, and large architectural or engineering projects are some of the relevant examples. We have also explored the implications of this observation to the issue of team formation, arguing that process-based teams are superior to function-based teams from the point of view of incentives. We have mentioned some empirical papers reporting evidence that seems consistent with our findings. However, our results have more concrete testable implications which we find challenging for further research. Is peer monitoring statistically more prevalent in organizations involving complementary tasks? Are process-based teams statistically more widespread than function-based teams? Are architectures of information, which are predicted to be more effective (in accordance with Proposition 3) are indeed more common? Both empirical and experimental methods seem adequate to address these and other related questions.

10 Appendix 1

Our comparison of complementarity and substitution (under PBE implementation) asserted that under substitution the cost of the optimal mechanism does not depend on the information structure (Proposition 2). In this Appendix we show by way of example that under substitution the principal can be strictly worse off with more peer information if there are multiple levels of effort and when the principal attempts to incentivize agents to exert less than full effort. We then use this framework to provide a two-agent result showing in general that the principal gains from more information among peers under complementarity and loses under substitution.

Consider a two-agent organization with each agent facing multiple levels of effort given by the sets $E_i = \{0, 1, 2, ..., k_i\}$ for $i = 1, 2$. A technology is now a function $p : E_1 \times E_2 \longrightarrow [0, 1]$. We say that p satisfies complementarity (substitution) if $p(e_1, e_2) - p(e_1, e_2 - 1)$ and $p(e_1 - 1, e_2) - p(e_1, e_2)$ are both increasing (weakly decreasing) in e_1 and in e_2. We retain the monotonicity of p (in both coordinates). Finally, we assume a constant cost c per unit of effort for the two agents.

Example:

Consider the following three-level example with $E_i = \{0, 1, 2\}$ given by the following technology:

$effort$	0,0	0,1	1,0	0,2	2,0	1,1	1,2	2,1	2,2
$prob$	0	0.5	0.6	0.85	0.7	0.9	0.95	0.95	0.98

It is easily verified that the above technology corresponds to substitution. Suppose that the principal wishes to sustain a success probability of at least 0.9 and seeks the optimal mechanism to achieve this goal. Clearly he will have to implement the effort levels (1,1). Standard arguments shown earlier imply that the optimal mechanism sustaining (1,1) in the simultaneous environment pays $\frac{c}{p(1,1)-p(0,1)} = \frac{c}{0.4}$ to agent 1 and $\frac{c}{p(1,1)-p(1,0)} = \frac{c}{0.3}$ to agent 2.

Consider now the sequential mechanism with agent 1 starting. To sustain the outcome $(1, 1)$ (as a subgame perfect equilibrium) we need to make $e_2 = 1$

a best response to $e_1 = 1$. So agent 2 should get $\frac{c}{p(1,1)-p(1,0)} = \frac{c}{0.3}$ as in the simultaneous game. But with this reward it is easy to check that agent 2's best response to $e_1 = 0$ will be $e_2 = 2$. Hence to guarantee that agent 1 chooses $e_1 = 1$ his reward needs to satisfy $vp(1,1) - c \leq vp(0,2)$. Hence, we need to pay agent 1 at least $\frac{c}{p(1,1)-p(0,2)} = \frac{c}{0.05} > \frac{c}{0.4}$, showing that the sequential game is more expensive for the principal.

The intuition behind the example is straightforward. The sequential environment increases the incentives of agent 1 to free ride knowing that agent 2 will increase effort if he/she observes agent 1 choosing $e_1 = 0$. The principal therefore has to increase the incentives of agent 1 so that he/she does not free ride. If the principal wishes to get full effort from both agents, then the cost of the mechanism in the simultaneous environment and the sequential one will be the same (as in Proposition 2). The reason is that under full effort and substitution players in the simultaneous game are already responding to beliefs under which they have the least incentive (i.e., the beliefs that the other player is exerting full effort) and the rewards accordingly need to be high to induce their effort.

Proposition 9: Assume p satisfies complementarity (substitution) and suppose that the principal wants to sustain a probability of success of at least α. Then the optimal simultaneous mechanism sustaining the probability of success α is strictly more expensive (weakly less expensive) than the corresponding sequential mechanism.

Proof: We start with the case of complementarity. Consider a pair of effort levels e_1, e_2 by players 1 and 2. Note that because of complementarity if an agent is better off deviating from (e_1, e_2) by reducing effort by a single level, then he would do even better by reducing his effort level to zero. The optimal simultaneous mechanism (v_1, v_2) implementing the effort levels (e_1, e_2) is therefore derived from solving the following incentive constraints: $v_1 p(e_1, e_2) - ce_1 = v_1 p(0, e_2)$ and $v_2 p(e_1, e_2) - ce_2 = v_1 p(e_1, 0)$. Hence, $v_1 = \frac{c}{p(e_1,e_2)-p(0,e_2)}$ and $v_2 = \frac{c}{p(e_1,e_2)-p(e_1,0)}$. Denote by $v(e_1, e_2)$ the total reward in the optimal mechanism implementing (e_1, e_2). To optimally implement a probability of success of at least α the principal has to pay $v(e_1^*, e_2^*)$, where $(e_1^*, e_2^*) = \arg\min\{v(e_1, e_2) | p(e_1, e_2) \geq \alpha\}$. We next show that for any pair of effort levels (e_1, e_2) the optimal sequential mechanism sustaining (e_1, e_2) pays the two agents a total payoff that is strictly less than $v(e_1, e_2)$. Indeed let $v_2^s = v_2 = \frac{c}{p(e_1,e_2)-p(e_1,0)}$ and $v_1^s = \frac{c}{p(e_1,e_2)-p(0,0)}$. It is easy to verify that (v_1^s, v_2^s) yields (e_1, e_2) as a unique subgame perfect equilibrium of the sequential game and because $v_2^s = v_2$ and $v_1^s > v_1$, the result is established for complementarity. Consider now the case of substitution. The relevant incentive constraints sustaining (e_1, e_2) under simultaneity are $v_1 p(e_1, e_2) - ce_1 = v_1 p(e_1 - 1, e_2) - c(e_1 - 1)$ and $v_2 p(e_1, e_2) - ce_2 = v_2 p(e_1, e_2 - 1) - c(e_2 - 1)$ yielding $v_1 = \frac{c}{p(e_1,e_2)-p(e_1-1,e_2)}$, and $v_2 = \frac{c}{p(e_1,e_2)-p(e_1,e_2-1)}$. As before under sequentiality the optimal mechanism (v_1^s, v_2^s) satisfies $v_2^s = v_2$. So it remains to compare v_1 and v_1^s. Suppose that agent 1 chooses $e_1 - 1$ in the sequential game and let $b(e_1 - 1)$ be agent 2's best response to $e_1 - 1$ in the sequential game. Because of substitution and because e_2 is a best response to e_1 we have $b(e_1 - 1) \geq e_2$. Hence if the optimal mechanism pays agent 1 v_1^s, it must be that $v_1^s p(e_1, e_2) - ce_1 \geq v_1^s p(e_1 - 1, b(e_1 - 1))$, yielding $v_1^s \geq \frac{c}{p(e_1,e_2)-p(e_1-1,b(e_1-1))}$. But because of monotonicity $p(e_1 - 1, b(e_1 - 1)) \geq p(e_1 - 1, e_2)$. Hence, $v_1^s \geq \frac{c}{p(e_1,e_2)-p(e_1-1,b(e_1-1))} \geq \frac{c}{p(e_1,e_2)-p(e_1-1,e_2)} = v_1$, which completes the proof. Q.E.D.

11 Appendix 2

Proof of Proposition 5: Fix probabilities q and q' with $q' > q$. Consider two organizations H and L in which peer information is given by the random graphs k_H and k_L with reliablity q and q' respectively. We can embed the two random graphs k_H and k_L in one sample space Θ with a joint probability distribution F such that the event "an arc exists from i to j in k_L" is a subset of the same event for the graph k_H, and such that F assigns these two events probabilities q and q' respectively.[7] Hence, denoting by $k_L(\theta), k_H(\theta)$ the corresponding deterministic graphs that form under the realization θ we have $k_L(\theta) \subset k_H(\theta)$ for all θ, with some θ for which this inclusion is strict. For an agent i and an order w we denote by $W(i)$ the set of agents following i including i himself, i.e., $W(i) = \{j \in N; w(j) \geq w(i)\}$.

For $x \in \{L, H\}$ denote by $k_x(\theta)|_{W(i)}$ the graph $k_x(\theta)$ restricted to the set of nodes $W(i)$; i.e., $k_x(\theta)|_{W(i)}$ is obtained from $k_x(\theta)$ by deleting all nodes in $N \backslash W(i)$ with all the arcs in and out of these nodes. $k_x(\theta)|_{W(i)}$ represents the monitoring opportunities among i's successors. For each j in $W(i)$ we denote by $j \rightarrow_x i$ if there exists a path $j = i_1, ..., i_r = i$ from j to i in the graph

[7] For example, if $q = 1/4$ and $q' = 1/2$ we can take the sample space to be the outcomes of two independent fair coin flips. Assign an arc from i to j in k_L iff both fall on tail and in k_H if at least one falls on tails. For the general construction see for example the textbook by Shaked and Shanthikumar (1994) (Theorem 1.A.1).

$k_x(\theta)|_{W(i)}$ such that $w(i_j) > w(i_{j+1})$. The intuition behind $j \to_x i$ is that there is a channel of messages from i to j based on the realized arcs such that the chain across which the messages pass involves only i's successors and is consistent with the order of moves. In other words, if $j \to_x i$, then agent j who succeeds i can indirectly monitor (through a sequence of other agents) whether i has exerted effort. We next denote by $S_x(i, w, \theta)$ the set of agents that can indirectly monitor i, i.e., $S_x(i, w, \theta) = \{j \in N; j \to_x i\}$. If all agents adopt the strategy by which they decide to shirk if and only if they (directly) detected the shirking of at least one agent, then the random set $S_x(i, w, \bullet)$ represents the set of agents that will be triggered to shirk as a consequence of i's shirking. For a technology p and given the strategies specified above the probability that the project will succeed following i's shirking (and assuming all his predecessors are exerting effort) is $p(N \backslash S_x(i, w, \bullet))$, which is a random variable. We will denote by $p_x(i)$ the expected value of this random variable.

We will now show that $p_H(i) > p_L(i)$. Indeed, we argued earlier that $k_L(\theta) \subset k_H(\theta)$ for all θ, with some θ for which this inclusion is strict. Therefore $k_L(\theta)|_{W(i)} \subset k_H(\theta)|_{W(i)}$ for all θ, with some θ for which this inclusion is strict, and hence $S_L(i, w, \theta) \subset S_H(i, w, \theta)$ for all θ, with some θ for which this inclusion is strict. Because p is strictly increasing we have $p(N \backslash S_H(i, w, \theta)) \geq p(N \backslash S_L(i, w, \theta))$ for all θ, with some θ for which this inequality is strict. Now $p_H(i) = \int_\theta p(N \backslash S_H(i, w, \theta))dF(\theta) > \int_\theta p(N \backslash S_L(i, w, \theta))dF(\theta) = p_L(i)$. We can now define $v_i^*(q) = c/[p(N) - p_L(i)]$ and $v_i^*(q') = c/[p(N) - p_H(i)]$. Since $p_H(i) > p_L(i)$ the proof will be completed if we show that $v_i^*(q)$ are the optimal INI mechanism under q (the same proof will apply for q').

Consider again the strategy profile defined earlier in the proof; i.e., an agent shirks if and only if he observes at least one of his predecessors shirking. We first show that this profile is a Nash equilibrium under the mechanism $v_i^*(q)$. By the definition of $p_q(i)$ and $v_i^*(q)$ we have $p(N)v_i^*(q) - c = v_i^*(q)p_i(q)$ and each player i is indifferent between investing and shirking if all other players adopt the strategy specified above. We now show that if some player i receives a reward $w_i < v_i^*(q)$ and all other players j receive $v_j^*(q)$ or more, then there exists no Nash equilibrium in which all players invest with probability 1. Assume by way of contradiction that such a Nash equilibrium exists. It must be the case that the strategy specifies that a player invests if he observes no shirking by others. Consider now player i's deviation in which he/she shirks regardless of the information he receives. Along the equilibrium path, if i shirks under the realization θ he will trigger at most the set of players $N \backslash S_L(i, w, \theta)$ to shirk (possibly a subset depending on the strategies of others), so with a reward $v_i^*(q)$ player i is either indifferent between investing and shirking or strictly prefers to shirk, and with a reward w_i he/she strictly prefers to shirk. To complete the proof we need to show that the strategy profile defined above is part of a PBE under the reward scheme that pays agent i $v_i^*(q)$ when the technology satisfies complementarity. For this we need to argue that a players who observes one of his peers shirking at an earlier stage would find it optimal to shirk as well. Supposing again that player i believes that all agents whom he/she cannot observe are exerting effort, let R be the set of players who acted before i and whom he/she observed shirking. We denote the $p^1_{q,R}$ (resp. $p^0_{q,R}$) to be the probability that player i assigns to the project succeeding under the distribution q when i exerts efforts (resp. shirks) and R. We will show that

$p^1_{q,R} - p^0_{q,R} < p(N) - p_q(i)$. (5.1)

For this we will make use of arguments we made use of in the proof of Proposition 4. Let θ be some realization of monitoring opportunity that player i can consider possible given the monitoring opportunities that transpired before i made his effort decision. For a set R of agents and the realization θ we denote by $R^*(\theta)$ the set of agents who see someone who sees someone ... that is in R under θ. Using the same argument as in the proof of Proposition 4 we can show that under complementarity of p we have

$p(N \backslash R^*(\theta)) - p(N \backslash (R \cup \{i\})^*(\theta)) \leq p(N) - p(N \backslash \{i\}^*(\theta))$ (5.2)

with strict inequality for some θs. But now (5.1) is obtained from (5.2) by integrating over the entire sample space to take expectations. It is now straightforward to show that given (5.1) and the fact that the mechanism is defined to be $v_i^*(q) = c/[p(N) - p_L(i)]$ agent i is better off shirking when observing the shirking of some of his predecessors. Q.E.D.

Proof of Proposition 6: We start with two lemmas:

Lemma 1: The optimal mechanism under the process-based structure pays the four agents the following rewards: $a_d : \frac{c}{p(N)-p(b_d,b_u)}$, $a_u : \frac{c}{p(N)-p(a_d,b_d,b_u)}$, $b_d : \frac{c}{p(N)-p(a_s,a_u)}$, $b_u : \frac{c}{p(N)-p(a_d,b_d,a_u)}$.

Lemma 2: The optimal mechanism under the function-based structure pays the four agents the following rewards: $a_d : \frac{c}{p(N)-p(b_d,a_u,b_u)}$, $a_u : \frac{c}{p(N)-p(a_d,b_d,b_u)}$, $b_d : \frac{c}{p(N)-p(a_d,a_u,b_u)}$, $b_u : \frac{c}{p(N)-p(a_d,b_d,a_u)}$

Proof of Lemma 1: Consider the following strategy profile in the game induced by the process-based structure: Players a_d and b_d exert effort, and player x_u exerts effort if and only if player x_d exerts effort, where $x \in \{a, b\}$. We first argue that this strategy profile is a PBE supported by beliefs that assign effort to every non-observable action. Assume first that a_d exerts effort. Then under this profile player a_u gets $p(N)v_{a_u} - c$ if he invests and $p(a_d, b_d, b_u)v_{a_u}$ if he shirks (where v_{a_u} is the reward specified in the lemma). Hence a_u is indifferent between his two actions when a_d invests. We now show that it is optimal for a_u to shirk if a_d shirks. If a_u invests after a_d shirks he receives $p(a_u, b_d, b_u)v_{a_u} - c$ and he gets $p(b_d, b_u)v_{a_u}$ if he shirks. But because a_u and a_d are complementary we have $p(N) - p(a_d, b_d, b_u) > p(a_u, b_d, b_u) - p(b_d, b_u)$ and therefore $p(a_u, b_d, b_u)v_{a_u} - c < p(b_d, b_u)v_{a_u}$. This implies that a_u is better off shirking. The same claims established above apply also with respect to b_d and b_u. To verify that the proposed strategy profile specifies optimal action for a_d and b_d we note that if, say, a_d shirks then a_u will shirk as well and the expected payoff for a_u is $p(b_d, b_u)v_{a_d}$, which is identical to $p(N)v_{a_d} - c$; hence it is optimal for a_d to exert effort and also for b_d to do so. We conclude that the strategy profile described above specifies optimal behavior for all agents both on and off the equilibrium path, and is therefore a PBE. To verify that the rewards specified in Proposition 6 are optimal consider a different reward vector for which some agent gets less and consider a strategy profile in which all agents exert effort. It is easy to see that such a profile cannot be an equilibrium. For d-type agents this is straightforward because they are already indifferent between shirking and investing under the rewards specified in the lemma. For u-type agents the argument is as follows. First, the strategy profile specified above cannot be an equilibrium because of the indifference. So consider for example the strategy for a_u in which he/she exerts effort regardless of the action taken by a_d. It can easily be shown that this strategy cannot be part of an equilibrium because of complementarity (we have shown that under the reward specified in Lemma 1 a_u is better off shirking if a_d shirks-all the more so for a lower reward). Hence, there exists no equilibrium in which all agents invest under the alternative reward scheme. Q.E.D.

Proof of Lemma 2: Consider the following strategy profile in the game induced by the function-based structure: All agents exert effort regardless of the information they obtain about the others. It is easy to show that given the rewards specified in Lemma 2 each agent is indifferent between exerting effort and shirking given that the rest are exerting effort. Hence the strategy profile above is a Nash equilibrium. To show that the profile specifies optimal action also off the equilibrium path we note that each agent who observes his peer shirking is weakly better off investing than shirking. Consider agent a_u who observed a_d shirking. If a_u invests his payoff is $p(a_u, b_d, b_u)v_{a_u} -$

c and it is $p(b_d, b_u)v_{a_u}$ if he shirks. But because of substitution we have $p(N) - p(a_d, b_d, b_u) \leq p(a_u, b_d, b_u) - p(b_d, b_u)$ and hence $p(a_u, b_d, b_u)v_{a_u} - c \geq p(b_d, b_u)v_{a_u}$, which establishes the claim. The proof that this mechanism is optimal is similar to that of in Lemma 1. Q.E.D.

We can now complete the proof of Proposition 9. Let v^1 and v^2 be the total reward paid by the principal under Lemmas 1 and 2 respectively. Then $v^1 - v^2 = \frac{c}{p(N)-p(b_d,b_u)} + \frac{c}{p(N)-p(a_s,a_u)} - \frac{c}{p(N)-p(a_d,b_d,b_u)} - \frac{c}{p(N)-p(a_d,a_u,b_u)}$. But because of strict monotonicity of p we have $p(N)-p(b_d, b_u) > p(N)-p(a_d, b_d, b_u)$ and $p(N) - p(a_s, a_u) > p(N) - p(a_d, a_u, b_u)$ or $\frac{c}{p(N)-p(b_d,b_u)} < \frac{c}{p(N)-p(a_d,b_d,b_u)}$ and $\frac{c}{p(N)-p(a_s,a_u)} < \frac{c}{p(N)-p(a_d,a_u,b_u)}$, implying that $v^1 - v^2 < 0$. Hence the optimal mechanism for the process-based structure is less expensive. Q.E.D.

Proof of Proposition 7: For each decision node x of m we denote by $P_m(x)$ the information set of m that contains the decision node x. We will now define by backward induction a reward scheme which we will then prove to be optimal with respect to the underlying information structure and technology.

We start with player n for which we define $v_n^* = \frac{c}{p(N)-p(N\backslash n)}$. For $n-1$ we define $v_{n-1}^* = \frac{c}{p(N)-p(N\backslash n-1)}$ if $\mathbf{1}\in P_n(1, 1, ...0)$ (which means that player n will not be informed of someone shirking after all players up to $n-1$ exerted effort and $n-1$ shirked). Otherwise we define $v_{n-1}^* = \frac{c}{p(N)-p(N\backslash\{n-1,n\})}$.

To define v_k^* we will define a set of players $S_k \subset \{k+1, ..., n\}$ by induction. Intuitively the set S_k denotes the set of player k's successors who will be triggered to shirk following k's shirking provided that all other $k's$ predecessors exerted effort. Specifically $k+1 \in S_k$ iff $\mathbf{1}\in P_{k+1}(1, 1, ...0)$. Suppose by way of induction that we have determined whether or not players $k+1, k+2, ..., k+j$ are in S_k. Then $k+j+1 \in S_k$ iff $\mathbf{1}\in P_{k+j+1}(1, 1, ..., 0, I_{k+1}, I_{k+2}, ..., I_{k+j})$, where $I_j = 0$ if $j \in S_k$ and $I_j = 1$ if $j \notin S_{k}$. Roughly speaking $k+j+1$ is in S_k iff this player is triggered to shirk by the shirking of player k (assuming that all preceding players shirk if and only if they observe others shirking players). With the above definition of S_k we can now specify the reward for player k which is $v_n^* = \frac{c}{p(N)-p(N\backslash(S_k\cup k))}$. We next show that the mechanism defined above is an optimal INI mechanism. Given $v_1, ..., v_n$ the following strategy profile is a Nash equilibrium: Each player exerts effort iff $\mathbf{1}= (1, 1, ..., 1)$ is in the information set he finds himself/herself at when called upon to make a decision. This follows from the following incentive constraint for player k:

$v_k^* p(N) - c = v_k^* p(N) - p(N\backslash(S_k \cup k))$ (*)

(i.e., k is indifferent between exerting effort and shirking given that the rest of the players are playing the profile described above). Furthermore the mechanism is optimal since any lower reward $v_k - \varepsilon$ for player k will result in shirking being a strict best response because of the constraint $(*)$. Finally, to show the monotonicity with respect to the information structure note that if I^1 is more transparent than I^2, then per definition the set S_k defined for the information structure I^2 is a subset of the set S_k defined for I^1. Given the definition of the rewards and because the technology p is monotonic this implies that $v^*(I^1, p) \leq v^*(I^2, p)$. Q.E.D.

Proof of Proposition 8: We define the mechanism in the same way as in the proof of Proposition 1. For this mechanism we have already shown that effort exertion by all players is a Nash equilibrium and that any mechanism that pays less to some agent will not admit full effort as a Nash equilibrium. To prove Proposition 8 we therefore need to show that the strategy profile specified in the proof of Proposition 7 sustains full effort as a Perfect Bayesian equilibrium when I satisfies *(LIS)* and when p satisfies complementarity. This profile instructs a player m to exert effort iff $\mathbf{1} \in P_m(x)$. The fact that these strategies are optimal on the equilibrium path (i.e., when $x = (1, ..., 1)$) has been proved in Proposition 7. We therefore have to show that these strategies prescribe a rational choice also off the equilibrium path. Put differently, we have to show that if a player can verify that shirking took place (i.e., when $\mathbf{1} \notin P_m(x)$) he should choose to shirk. To this end we will compare the consequences of player m's decision to shirk and to exert effort under two different type of scenarios: (1) when all his predecessors exerted effort and (2) when some of his predecessors shirked. We define by S^1 the set of agents who exert effort in the continuation of the game following the path in which all of m's predecessors including m himself exerted effort (i.e., following the path $(1, ..., 1)$). We denote by S^0 the same set but now following the path in which $m's$ predecessors exerted effort but m himself decided to shirk (i.e., following the path $(1, 1, ..., 1, 0)$). Now consider a different path denoted by x involving $m's$ predecessors (in which some of them are shirking) and define $\widehat{S^1}$ and $\widehat{S^0}$ but now with respect to the paths $(x, 1)$ and $(x, 0)$ respectively. In the four cases we are assuming that players acting after m use the strategy "exert effort iff $\mathbf{1} \in P_j(x_j)$." Per definition we have the following inclusions: $S^0 \subset S^1$, $\widehat{S^0} \subset \widehat{S^1}$, $\widehat{S^1} \subset S^1$, and $\widehat{S^0} \subset S^0$. Hence $\widehat{S^0} \subset S^0 \cap \widehat{S^1}$. We will now show that by the property of *(LIS)* we have $\widehat{S^0} = S^0 \cap \widehat{S^1}$. This will be shown by induction on the order of the players starting with player $m+1$. Suppose that $m+1 \in S^0$. Then $m+1$ chooses effort following the path $(1, 1...1, 0)$. If

$m+1 \in \widehat{S^1}$. Then $m+1$ chooses effort following the path $(x,1)$. By (LIS) $m+1$ must also choose effort under $(x,0)$ and hence $m+1 \in \widehat{S^0}$. Assume now by induction that for each player j between $m+1$ and $m+k$ we have that $j \in S^0 \cap \widehat{S^1}$, which implies $j \in \widehat{S^0}$. Consider player $m+k+1$. Let z_1, z_2, z_3 be the paths of length $m+k$ that evolve from the paths $(1,1...1,0)$, $(x,1)$ and $(x,0)$ respectively. By the induction $z_3 = z_1 \wedge z_2$, and the claim for $m+k$ follows now directly from (LIS). Consider now a technology p with increasing returns to scale. As a function $p: 2^N \to [0,1]$ the technology p satisfies the following condition: If $T_1 \subset T_2$ and Q_1, Q_2, R_1, R_2 are sets whose intersection with T_2 is empty and such that $Q_1 \subset Q_2, R_1 \subset R_2$, $R_1 \subset Q_1, R_2 \subset Q_2$ and $R_2 \backslash R_1 \subset Q_2 \backslash Q_1$, then $p(T_2 \cup Q_2) - p(T_2 \cup Q_1) > p(T_1 \cup R_2) - p(T_1 \cup R_1)$. This is shown by using the definition of convexity and applying it player by player (it also follows from the following claim: for a convex function f, if $x < y < w < z$ and $w - x < z - y$, then $f(z) - f(y) > f(w) - f(x)$).

$p(\{1,...,m\} \cup S^1) - p(\{1,...m-1\} \cup S^0) > p(\{j \leq m-1 | x_j = 1\} \cup \{i\} \cup \widehat{S^1}) - p(\{j \leq m-1 | x_j = 1\} \cup \widehat{S^0})$. (8.1)

This is shown by setting $T_1 = \{j \leq m-1 | x_j = 1\}, T_2 = \{1,2,...,m-1\}, Q_1 = S^0, Q_2 = S^1 \cup \{m\}, R_1 = \widehat{S^0}$ and $R_2 = \widehat{S^1} \cup \{m\}$, and using the fact that $\widehat{S^0} = S^0 \cap \widehat{S^1}$, which we established earlier. Note now that the reward for player m as defined in the proof of Proposition 7 can also be written as $v_m = \frac{c}{p(\{1,...,m\} \cup S^1) - p(\{1,...,m-1\} \cup S^0)}$. To show that the strategy profile defined in Proposition 7 is PBE we have to show that player m is better off shirking if he can verify that some of his predecessors shirked. Let x such path, i.e., $\mathbf{1} \notin P_m(x)$. If m exerts effort his expected payoff will be $v_m p(\{j \leq m-1 | x_j = 1\} \cup \{i\} \cup \widehat{S^1}) - c$. If instead he shirks his expected payoff is $v_m p(\{j \leq m-1 | x_j = 1\} \cup \widehat{S^0})$. Using the inequality (8.1) we establish that the latter is greater. Hence, player m is better off shirking following a path in which $\mathbf{1} \notin P_m(x)$, which completes the proof. Q.E.D.

References

[1] Babaioff, M., M. Feldman and N. Nisan (2006) "Combinatorial Agency," Mimeo, UC Berkeley.

[2] Babaioff, M., M. Feldman and N. Nisan (2006) "The Price of Purity and Free-Labor in Combinatorial Agency," Mimeo, UC Berkeley.

[3] Baliga, S. and T. Sjoestrom (1998) "Decentralization and Collusion," *Journal of Economic Theory* 83, 162-232.

[4] Che, Y.K. and S. W. Yoo (2001) "Optimal Incentives for Teams," *American Economic Review* 91, 525-541.

[5] Gould, E. D. and E. Winter (2005) "Rational Peer Effects between Workers: Evidence from Professional Baseball," Mimeo, The Hebrew University of Jerusalem.

[6] Heywood, J.S. and U. Jirjahn (2004), "Teams, Teamwork and Absence," *Scandinavian Journal of Economics* 106, 765-782.

[7] Holmstrom, B. (1982) "Moral Hazard in Teams," *Bell Journal of Economics* 13, 324-340.

[8] Holmstrom, B. and P.Milgrom (1990) "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics*, 146, 85-105.

[9] Ichino, A. and G. Maggi (2000) "Work Environment and Individual Background Explaining Shirking Differentials in Large Italian Firms," *Quarterly Journal of Economics* 115, 1057-1090.

[10] Itoh, H. (1991) "Incentives to Help Multi-Agent Situations," *Econometrica* 59, 611-636.

[11] Jackson, M.(2006) "A Survey of Models of Network Formation: Stability and Efficiency," Group Formation in Economics; Networks, Clubs and Coalitions, edited by Gabielle Demange and Myrna Wooders, Cambridge University Press.

[12] Kandel, E. and E. Lazear (1992) "Peer Pressure and Partnerships," *Journal of Political Economy* 100, 801-817.

[13] Knez, M. and D. Simester (2001), "Firm-Wide Incentives and Mutual Monitoring at Continental Airlines," *Journal of Labor Economics* 19, 743-772.

[14] Moretti, E. and A. Mas (2006) "Peers at Work," mimeo, UC Berkeley and NBER.

[15] Roth, A., O. Fujiwara, V. Prashnikar and S. Zamir (1991) "Bargaining and Market Behavior in Jerusalem, Ljubliana, Pittsburgh and Tokyo: An Experimental Study," *American Economic Review* 73, 362-72.

[16] Shaked, M. and J.G. Shanthikumar (1994) "Stochastic Orders and their Applications," Academic Press, San Diego CA.

[17] Segal, I. (2003) "Collusion, Exclusion, and Inclusion in Random-Order Bargaining," *Review of Economic Studies* 70, 439-460.

[18] Teasley, S. D., L. Covi, M.S. Krishnan, and J. S. Olson, (2002) "Rapid Software Development through Team Co-location." *IEEE Transactions on Software Engineering* 28, 671-683.

[19] Winter, E. (2004) "Incentives and Discrimination," *American Economic Review*, 94, 764-773.

[20] Winter, E. (2006) "Optimal Incentives for Sequential Production Processes, " *Rand Journal of Economics* (forthcoming).

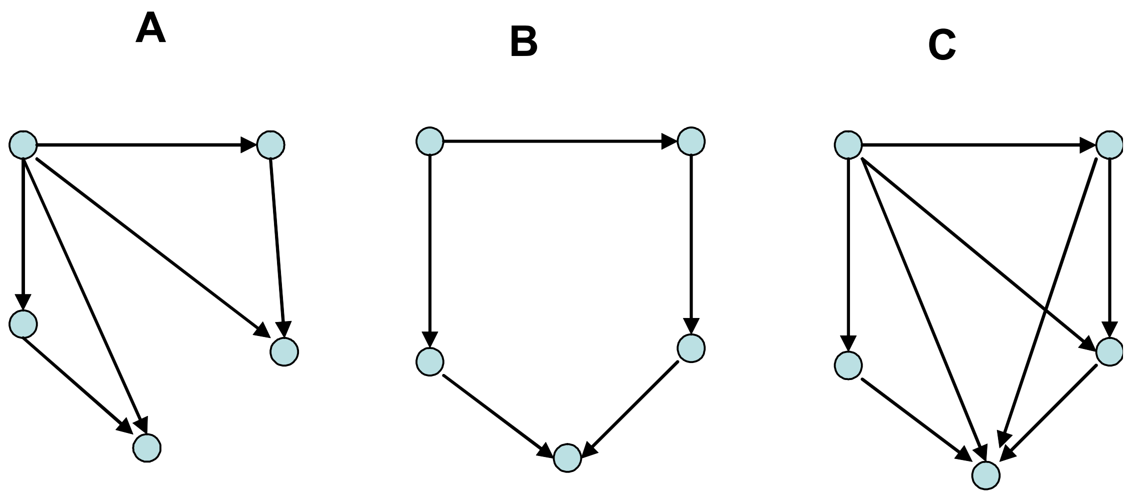

Figure 1: