On Transparency in Organizations*

Philippe Jehiel†

28th November 2011

Abstract

Should workers be informed of the incentive schemes governing their team co-workers? Should workers be told on which performance measure they will be assessed? Should workers be told how important the tasks are for the organization? This paper shows in an abstract moral hazard setup why in rich enough environments it is always strictly desirable that some aspects of the interaction be kept unknown to the workers. Thus, full transparency is argued to be suboptimal quite generally.

1 Introduction

A central question of organizational design is about how to make workers exert more productive effort. Classic economic analysis requires using instruments such as wages and bonuses indexed on what is observable to the principal so as to best align workers' interest with the organizational objective (see Holmström (1979-1982) or Myerson (1982) for classic references). In this paper, I ask myself whether it can be beneficial for the organization that the workers be incompletely informed of the details of the interaction

*I would like to thank seminar participants at Cesifo IO workshop, Paris 1, the workshop on multi-dimensional mechanism design in Bonn, Warwick workshop, StonyBrook game theory conference, Queen Mary University workshop, the conference on gaming in Bonn and in particular David Martimort, Paul Milgrom, Andy Newman, and Larry Samuelson for insightful comments. I also thank the European Research Council for financial support.

†PSE, 48 boulevard Jourdan, 75014 Paris, France and University College London ; jehiel@enpc.fr

they are in. It should be noted that I do not necessarily insist that instruments such as wages and bonuses be optimally adjusted,[1] but I take the view that how information is distributed in the organization can (at least partly) be controlled by the organization.

From another perspective, I am interested in whether it is a good idea for organizations to be as transparent as possible by which I mean to let workers know everything the organization knows including the performance measures that are available in the organization, the compensation schemes applying to co-team workers, how important the tasks are for the organization etc.

I use an abstract setup to address this question. Agents can potentially interact over a family of moral hazard problems which are parameterized by a state variable α that takes its value in an s-dimensional space. In each moral hazard problem α, agent i has to choose an unobservable action a_i in an n_i-dimensional space. In some interactions there may be only one agent, in others there may be several agents. Monetary instruments may potentially be used in which case they must be indexed on what is observed by the principal and verifiable by third parties.

Such an abstract framework is extremely general allowing me to speak of moral hazard in team interactions (Holmström, 1982), multi-tasking (Holmström and Milgrom, 1991; see also Baker, 2000), delegation and authority (Aghion and Tirole, 1997), Sender-receiver interactions (Crawford and Sobel, 1982) to name just a few (classic) applications.

I address the transparency question by comparing the organizational payoff when agents have full information about α and when (at least) one agent has incomplete information about α where I allow myself to choose the information structure of the agents freely. When there exists a coarse information structure that allows the organization to strictly enhance its objective, I say that non-transparency is desirable.

The main result of this paper is that when the space of moral hazard problems is rich enough (i.e., α has dimension strictly larger than the dimensionality of at least one agent's action space), non-transparency is desirable. More generally, assuming that only partial aspects of α can be kept secret to agents, the non-transparency result generically holds, as soon as the dimension of α that can be hidden exceeds the dimension of the action space of agents. Importantly, the conclusion need not be the same if the dimensionality

[1] There may be other constraints limiting the scope for optimal adjustments.

of the space of moral hazard problems is the same as or lower than the dimensionality of agents' action spaces, as I demonstrate through examples.

In my view, such a result is significant because in most cases the space of moral hazard problems would typically vary over more dimensions than there are dimensions in agents' action spaces. This is because parameters of the moral hazard problem include (among other things) how important the tasks are to the organization and how costly it is to change the action in each dimension of the action space of each agent (and the organization can a priori hide to a worker the cost structure of co-workers as well the magnitude of the stakes for the organization).[2] Thus, the conclusion established in this paper about the desirability of non-transparency should be expected to apply to quite generally.

The basic trade-off concerning the effect of coarsening an agent's information about α is as follows. If an agent receives incomplete information about α, it makes it easier to satisfy the incentive constraints (i.e., to make the agent choose what he is supposed to choose) as compared with the complete information case. This is because incomplete information allows the principal to aggregate the various incentive constraints into a single one, simpler to satisfy. But, incomplete information also forces the agent not to be able to adjust his action to the problem α, which may sometimes be undesirable. In general, the comparison between the complete and incomplete information case can go either way. But, when the state variable α varies over more dimensions than there are dimensions in the action space of one of the agents, one can always find an information structure of this agent such that the positive effect dominates the negative one, thereby showing that some form of non-transparency is always optimal.

To suggest the importance of the dimensionality requirement for the result, consider a family of moral hazard in team interactions which would vary only according to the degree of complementarity between the team members' efforts. Hiding the degree of complementarity to workers would not in general be a good idea because it would average the effort level across the various team interactions whereas it would typically be preferable to induce more effort when the degree of complementarity is higher and less effort when it is lower. However, if the team interactions differ also in how important the tasks are for the

[2] Alternatively, the organization can ensure that the bonus schemes of agents (that typically depend on the realization of the various dimensions of effort) is not publicly available.

organization, not letting the workers know whether the interaction is of little importance for the organization but with high degree of complementarity or of great importance to the organization but with low degree of complementarity may be desirable because it allows the organization to take advantage of the easing of the incentive constraints brought by the high degree of complementarity in the less important interaction for the task that is more important for the organization.

From a more theoretical perspective, when the state variable α varies over more dimensions than there are dimensions in the action space of one of the agents, simple topological arguments reveal that there must be a manifold of positive dimension in the space of α in which were the agent to be fully informed of α he would play in the same way over the various α in the manifold. Clearly, if the agent were only informed that α lies in the manifold rather than being informed of the exact realization of α, this would make no difference. The idea behind the transparency result is to bundle α_1 and α_2 in an information set where α_1 lies in the manifold and α_2 lies in the neighborhood of the manifold (away from α_1, which is possible because the manifold has a strictly positive dimension). For generic objective functions, the main result of this paper shows that one can always find such an information structure with the effect of strictly enhancing the organizational objective as compared with the complete information case.

Related literature:

1) In a framework allowing for both private information and private actions unobservable to the principal, Myerson (1982) considers the optimal mechanism and shows that the principal can restrict herself to incentive-compatible direct coordination mechanisms in which agents report their information to the principal who then recommends to them decisions forming a correlated equilibrium.[3] In the context of the present analysis, this implies that if a mediator could be used to make recommendations to agents in the various α, the optimal mechanism (à la Myerson) could be implemented with agent(s) being only informed of which action to play and not which α prevails.

The non-transparency result of this paper is of a different nature. First, it applies even outside the context of optimal mechanisms (including situations in which no mediator can

[3] Such an observation can be viewed as an expression of the revelation principle in such settings. See Rahman (2010) for an interesting recent application of Myerson's insight.

be used or monetary instruments are limited or even inexistent). Second, Myerson's observation does not imply that providing information about α would be strictly detrimental to the principal, which should be contrasted with our main finding that some form of incomplete information is strictly desirable relative to the complete information case.

2) The question addressed in this paper is related to the optimal information disclosure question addressed in Rayo and Segal (2010) and to the Bayesian persuasion question addressed in Kamenica and Gentzkow (2011). These authors consider Sender-Receiver interactions in which the Sender possesses private information and the Receiver chooses an action based on the information he has (or infers from the Sender's communication). And they ask themselves which disclosure policy the Sender should commit to so as to maximize her expected payoff. Restricting attention to Principal-agent interactions with no monetary transfers, my non-transparency result says that a Sender whose private information has higher dimension than the dimension of the action of the Receiver can do better than disclosing everything he knows. It turns out that an illustration of the non-transparency result can be found in the context studied by Rayo and Segal in which the Sender possesses information both on how valuable to the Sender and the Receiver the project is, and the Receiver chooses a probability of choosing the project (a one-dimensional action). Rayo and Segal' setup satisfies the dimensionality property required for the non-transparency result, and indeed Rayo and Segal observe that the best disclosure policy is never to disclose everything to the Receiver.[4] I note that Kamenica and Gentzkow's investigation of Bayesian persuasion focuses on when it is best for the Sender to disclose some information (as opposed to none) to the Sender, which is at the other extreme of the non-transparency question asking when it is best not to disclose as much as possible.

3) From a more general perspective, the idea that hiding some information might help achieving better outcomes has been considered in a number of settings (Abreu-Milgrom-Pearce (1991), Lizzeri-Meyer-Persico (2002) or Aoyagi (2010) to name just a

[4] Rayo and Segal characterize fully the optimal disclosure policy in their more structured steup. By contrast, this paper shows how general the non-transparency result is when the dimensionality property holds without characterizing the optimal disclosure policy.

As Rayo and Segal note, almost no paper in the Sender-Receiver tradition has considered the case in which the Sender possesses a multi-dimensional private information. Rayo and Segal together with this paper (if particularized to the Sender-Receiver case) are the exceptions.

few). The logic of the results reported in these papers is related to the basic observation that coarsening the information of agents helps alleviating their incentive constraints. What the main Theorem of this paper shows is that the strict desirability of having incomplete information applies quite generally, whenever the state space is of a sufficiently large dimensionality. None of the above mentioned papers has highlighted the role of dimensionality.[5]

4) The desirability of transparency addressed in this paper is also related to the theme of the value of information in strategic interactions (see Hirshleifer (1971) and Bassan et al (2003) for, respectively, a pioneering and more recent contribution on the subject). From a more applied perspective, Milgrom and Weber (1982) show in the context of auctions with affiliated signals that providing more information to bidders increases revenues (an insight known as the linkage principle which should be contrasted with our non-transparency result, see also Ottaviani and Prat (2001) for a similar investigation for a price discriminating monopolist).

5) Finally, it should be mentioned that there have been many other approaches to transparency in organizations. I mention here just a few to help locating the present contribution in the literature. In seminal contributions, Holmström (1979-1982) has shown in static moral hazard problems that it is always best that the principal be as informed as possible, as it allows her to better monitor the agent(s). In subsequent works, (Crémer (1995), Dewatripont et al. (1999) or Prat (2005) to name just a few), dynamic considerations have been introduced, imposing some limited commitment capabilities on the Principal's side. There, less information for the Principal may help the Principal, as it may alleviate her commitment problems. Note that this line of research is more concerned with changing the information held by the Principal whereas the focus of this paper is on the information held by the agents (as well as the feedback transmitted to them).[6]

The rest of the paper is structured as follows. Preliminary examples are provided in

[5] While Battaglini (2006) emphasizes the role of dimensionality in a multi-agent moral hazard problem, the question addressed in his paper is of a different nature (i.e., it is about how the dimensionality of the signal space helps distinguishing deviators from non-deviators).

[6] Other approaches to transparency in which the information of third party is considered includes Gavazza and Lizzeri (2009) and Dubey and Geanakoplos (2010) among others. Winter (2010) considers a different form of transparency, i.e. whether it is good or bad for the organization to let agents observe their peer's effort choices, and he asks himself which observability structure allows the organization to best rule out undesirable outcomes in a team problem.

Section 2. A general framework is presented in Section 3. Our two main questions are formally stated in Section 4. Section 5 contains the main results as well as a discussion of these. Section 6 illustrates the key role played by the dimensionality in deriving the main insights. Concluding remarks appear in Section 7.

2 Preliminary examples

In this Section I provide simple examples illustrating that some form of non-transparency may sometimes be good for the organization. In the rest of the paper, I will develop an abstract setup in which the questions will be addressed with much greater generality.

2.1 A Principal-Agent example

Consider two one-agent moral hazard problems α_x, α_y in which the agent must perform two tasks $a_x, a_y \in [0, 1]$. The cost incurred by the agent takes the additive form $c(a_x, a_y) = h(a_x) + h(a_y)$ both in α_x and α_y where $h(0) = h'(0) = 0$ and $h(\cdot)$ is assumed to be increasing.

The output exhibits complementarities between the two tasks. Specifically, output is given by $z = a_x a_y + \varepsilon$ where ε is the realization of some normal distribution centered around 0.

Output is not assumed to be observable (at least within a reasonable amount of time). In situation α_x, only $q_x = a_x + \varepsilon_x$ is observed by the principal (and verifiable by third parties) and similarly in situation α_y, only $q_y = a_y + \varepsilon_y$ is observed by the principal where ε_x and ε_y are the realizations of independent normal distributions centered around 0. Situations α_x and α_y are assumed to be equally likely in the organization.

I assume that wages must be non-negative. The principal's instrument thus boils down to offering bonus schemes $w_x(q_x) \geq 0$ in α_x or $w_y(q_y) \geq 0$ in α_y. The agent and the principal are assumed to be risk neutral. The agent gets a payoff equal to $w - c(a_x, a_y)$ when he earns w and exerts effort $a = (a_x, a_y)$; the principal gets an expected payoff equal to $a_x a_y - w$ under the same circumstances.

It is rather easy to see the advantage of not letting the agent know whether α_x or α_y. To see this, assume first that the agent knows α_x. Then clearly, the agent will pick $a_y = 0$

whatever $w_x(\cdot)$ (this is because a_y does not affect q_x and any $a_y > 0$ would induce strictly positive extra cost). Thus, expected output is 0 in the full information case (and $w_x(\cdot)$ and $w_y(\cdot)$ are optimally set at 0).

By contrast, consider the case in which the agent does not know whether α_x or α_y at the time he must choose a_x, a_y.[7] It is fairly easy to induce $a_x > 0$ and $a_y > 0$ through the choice of strictly increasing $w_x(\cdot), w_y(\cdot)$ because now the agent chooses (a_x, a_y) so as to maximize:

$$\frac{1}{2}E(w_x(a_x + \varepsilon_x)) + \frac{1}{2}E(w_y(a_y + \varepsilon_y)) - c(a_x, a_y).$$

More precisely, one can establish that the full information benchmark is dominated by the coarse information case whenever $h''(0) < \frac{1}{2}$ (by considering schemes of the form $w_z(q_z) = \max(0, \omega q_z)$ for sufficiently small ω and $z = x, y$).

In more intuitive terms, not letting the agent know whether α_x or α_y makes it easier to let the agent exert effort on both tasks because he does not know which one will be used as a performance measure to reward him. By contrast, when the agent knows that he will be assessed only on the basis of a_x (which is a consequence of the monitoring technology in α_x) he has no incentive to exert effort on a_y, which when the two tasks are sufficiently complement, is very detrimental to the output.[8] A related intuition appears in a recent paper by Ederer et al. (2008) who consider mixed moral hazard Principal-agent problems in which the agent has superior information.

Of course, the above example should not be interpreted to mean that coarse information is always good. An obvious potential disadvantage of coarse information is that the agent can no longer adjust his effort decision to the exact conditions governing the moral hazard interaction. In general, coarse information has the advantage of easing the incentive constraints (because it aggregates several incentive constraints into a single one, thus easier to satisfy), and it has the disadvantage of making the strategy less sensitive to the environment (the strategy must be measurable with respect to a coarser information

[7] After a_x and a_y are chosen, the agent is informed whether $\alpha = \alpha_x$ or α_y so that the bonus scheme can be implemented.

[8] One might alternatively assume that there are two self-interested agents, one in charge of choosing a_x (with cost $c(a_x)$) and one in charge of choosing a_y (with cost $c(a_x)$) and not having agents know whether $\alpha = \alpha_x$ or α_y is preferable when the two tasks are sufficiently complement.

partition). The trade-off between these two forces can go either way in general. But, as will be seen later on, when the space of problems is rich enough, one can always find a coarse information structure that strictly enhances the designer's objective as compared with the full information benchmark.

2.2 A moral hazard in team example

Consider the following two moral hazard in team interactions denoted X and Y, which are equally likely. In each interaction, two agents $i = 1, 2$ must simultaneously exert an effort $a_i \in \mathbb{R}^+$. The outcome of the team interaction is either successful with probability $p(a_1, a_2; \beta)$ or it is not successful with probability $1 - p(a_1, a_2; \beta)$ where

$$p(a_1, a_2; \beta) = a_1 + a_2 + \beta a_1 a_2.$$

The parameter $\beta > 0$ reflects the complementarity between agents' efforts. The reward to the organization in case of success depends on the interaction. It is R_X in interaction X and R_Y in interaction Y. The bonus received by the agents in case of success is set independently of the interaction. To simplify the exposition, this bonus is null for agent 2 and it is w for agent 1. Agent 2 thus exerts effort up to the point the effort becomes costly. The effort cost to agent 2 is assumed to depend on the interaction so that agent 2 exerts no effort in interaction Y and effort $\overline{a} > 0$ in interaction X. The effort cost to agent 1 is $\frac{1}{2}(a_1)^2$ in both X and Y.

Our question of interest is whether agent 1 should be informed whether the interaction is X or Y. Agent 2 exerts more effort in X than in Y, and there is complementarity in effort in both X and Y. So if Y is sufficiently more important than Y to the organization ($R_Y > R_X$), not letting agent 1 know whether the interaction is X or Y may be beneficial as it may allow to boost agent 1's incentive scheme in the more important task Y thanks to the higher effort level of agent 2 in task X.

Formally, if agent 1 knows he is in X, he exerts effort a_1 so as to maximize $(a_1 + \overline{a} + \beta a_1 \overline{a})w - \frac{1}{2}(a_1)^2$. That is, $a_1^X = (1 + \beta\overline{a})w$.

If agent 1 knows he is in Y, he exerts effort a_1 so as to maximize $a_1 w - \frac{1}{2}(a_1)^2$. That is, $a_1^Y = w$.

So in the full information case, the expected organizational payoff writes:

$$\pi^{Full} = \frac{1}{2}[(1+\beta\overline{a})^2 w + \overline{a}](R_X - w) + \frac{1}{2}w(R_Y - w).$$

If agent 1 does not whether X or Y, he chooses a_1 so as to maximize $(a_1 + \frac{\overline{a}}{2} + \beta a_1 \frac{\overline{a}}{2})w - \frac{1}{2}(a_1)^2$. That is, $a_1^C = (1+\beta\frac{\overline{a}}{2})w$, and the expected organizational payoff writes:

$$\pi^{Coarse} = \frac{1}{2}[(1+\beta\overline{a})(1+\beta\frac{\overline{a}}{2})w + \overline{a}](R_X - w) + \frac{1}{2}(1+\beta\frac{\overline{a}}{2})w(R_Y - w).$$

Thus,

$$\pi^{Coarse} - \pi^{Full} = \frac{\beta\overline{a}w}{4}[(R_Y - w) - (1+\beta\overline{a})(R_X - w)]$$

which is positive if R_Y is sufficiently larger than R_X.

3 A general framework

Throughout the paper, moral hazard problems with one or two agents parameterized by $\alpha \in \mathbb{R}^s$ are considered. The parameter α is assumed to be distributed according to a smooth (i.e., continuously differentiable) density $p(\alpha)$ that is strictly positive on some open subset of $\mathbb{R}^s$. Extensions to more than two agents raise no difficulties. In every moral hazard problem α, agent $i = 1, 2$ chooses an action a_i in A_i, an open subset of $\mathbb{R}^{n_i}$. Agents choose their actions simultaneously, that is, without observing the actions chosen by the other agent.[9]

While the designer is assumed to know α, I consider various informational assumptions regarding what the agents know about α. In addition, the designer may or may not (depending on the application) be allowed to use instruments $w = (w_1, w_2) \in \mathbb{R}^{\omega_1} \times \mathbb{R}^{\omega_2}$ that affect agents 1 and 2' incentives respectively, and that are based on what can be observed by the designer and third parties (typically actions a_i are not observable by the designer or they are not verifiable by third parties to make the problem non-trivial).

In problem α, agent i's expected payoff is $u_i(a_1, a_2; w, \alpha)$. The designer's expected payoff is $\pi(a_1, a_2; w, \alpha)$. The relevant range of a_i, w, α will assumed to be bounded

[9] Extensions to the case of sequential play would raise no difficulty.

throughout the paper.

It should be mentioned that in the above formulation, agents' participation constraints are not explicitly taken into account. Yet, when one of the actions in A_i ensures that agent i gets at least what he can get outside the interaction (whatever a_j), then agent i's participation constraint is automatically satisfied. Participation constraints will be further discussed later on after our main result is stated. Mechanisms allowing the use of mediators (à la Myerson (1982)) will also be discussed then.

The framework covers lots of classic moral hazard problems. To mention, just a few:

Multi-task and moral hazard **(Holmström and Milgrom, 1991)**

Even though the general framework admits several agents, it may obviously be particularized to one agent moral hazard problems (simply by freezing one of the two agents). Given that no restrictions are being made on the dimensionality of the action space of the agent, the framework covers the important application of multi-tasking (such as considered in subsection 2.1). For example, a single agent may consider exerting effort a_x, a_y in two tasks x and y with a corresponding cost $g(a_x, a_y)$. The expected output $z = h(a_x, a_y)$ is a function of the effort produced in the two dimensions, and the designer only observes some signal $q = r(a_x, a_y) + \varepsilon$ where ε is the realization of a normal distribution with variance σ^2 and mean 0. The designer may use a signal-dependent wage schedule $w(q)$ as instrument. The objective of the designer assumed to be risk neutral writes $z - E(w(q))$ and the agent assumed to exhibit constant absolute risk aversion gets an expected utility: $-E\exp[-\rho(w(q) - g(a_x, a_y)]$. The multi-task problem is parameterized by $\alpha = (h, r, \sigma, \rho, g)$.[10]

Moral hazard in teams **(à la Holmström, 1982)**

Consider a class of problems such as the one introduced in Subsection 2.2. Two risk-neutral agents 1 and 2 in a team simultaneously exert effort a_1 and a_2 say within the range $[\underline{a}, \overline{a}]$. With probability $\widetilde{p}(a_1, a_2; \beta) = a_1 + a_2 + \beta a_1 a_2$ the team is successful giving reward R

[10]In order to ensure that α lives in a space of finite dimension as assumed above, one could further parameterize the cost function $g(\cdot)$ and the function $r(\cdot)$ so that they depend on a finite number of parameters.

to the organization where the parameter $\beta \in [\underline{\beta}, \overline{\beta}]$ reflects the degree of complementarity between the effort levels chosen by the two agents.[11]

Efforts are not directly observable, only success is. Agents must receive non-negative wages in all events. The instruments available to the designer are the bonuses w_1 and w_2 given to agents 1 and 2 respectively in case of success (the wage in case of failure is optimally set at 0). Letting $g_i(a_i)$ denote the cost to agent i of making effort a_i, this moral hazard in team problem falls in the general framework just defined with:

$$\begin{aligned} u_i(a_1, a_2; w, \alpha) &= \widetilde{p}(a_1, a_2; \beta) w_i - g_i(a_i) \\ \pi(a_1, a_2; w, \alpha) &= \widetilde{p}(a_1, a_2; \beta)(R - w_1 - w_2) \end{aligned}$$

Here, the team problem is parameterized by $\alpha = (\beta, R, g_1, g_2)$, the profile of complementarity, reward and cost parameters.

Models of authority **(Aghion and Tirole, 1997)**

Agent 1 exerts effort a_1 to find out which project to adopt. The principal, agent 2, can exert effort so as to improve upon the choice of the agent. A good project for the agent gives him a private benefit b and a good project for the principal gives her a private benefit B. The probability that a good project for the agent (agent 1) is also a good project for the principal (agent 2) is γ and the probability that a good project for the principal is a good project for the agent is β. Identifying the effort levels with the probability of finding a good project and letting $g_1(a_1)$ and $g_2(a_2)$ denote the costs of efforts made by the agent and the principal, respectively, the expected utilities of the agent and the principal write:

$u_1(a_1, a_2; \alpha) = a_2 \beta b + (1 - a_2) a_1 b - g_1(a_1)$

$u_2(a_1, a_2; \alpha) = a_2 B + (1 - a_2) a_1 \gamma b - g_2(a_2)$

Here the designer's objective coincides with 2's objective $\pi(a_1, a_2; \alpha) = u_2(a_1, a_2; \alpha)$, and the authority problem is parameterized by $\alpha = (\beta, \gamma, b, B, g_1, g_2)$, the profile of congruence, private benefit and cost parameters.[12]

[11] We should assume that $2\overline{a} + \overline{\beta}\overline{a} < 1$ so that $p(a_1, a_2; \beta) \in (0, 1)$ for all a_1, a_2 in $[\underline{a}, \overline{a}]$.

[12] In Aghion-Tirole's model, there are no monetary instruments.

Sender-Receiver interaction **(Crawford-Sobel 1982)**

In such setups, the Sender is informed of the state of the economy ω and the Receiver must choose an action a. The utility of the Receiver and the Sender depend on the state ω and the action: $u_R(a;\omega)$ and $u_S(a;\omega)$. Such a setting obviously falls in the general class of problems studied in this paper by identifying ω with α.

4 Transparency

Within the framework described in Section 3, I ask the following question.

Question. Can it be beneficial for the designer that at least one agent, say agent 1, be partially rather than fully informed of α?

When the answer to the above question is affirmative, I say that some form of non-transparency is desirable.

The above transparency question echoes familiar investigations in economic theory. For example, it is similar to a question addressed by Milgrom and Weber (1982) in standard auctions with affiliated information. There, in a context of one-dimensional adverse selection auction models, Milgrom and Weber show that under affiliation, it is optimal for the seller to release as much information as she can to the bidders (see also Ottaviani and Prat (2001) for a similar investigation in the monopoly problem).

In the context of moral hazard problems, the above question has received less attention. Of course, the above question is related to the issue of optimal information disclosure as studied in a Sender-Receiver setup by Rayo and Segal (2010) or to the issue of Bayesian persuasion as studied (also in a Sender-Receiver setup) by Kamenica and Gentzkow (2011). Compared to Rayo and Segal, the question studied here is more limited (since there is no characterization of the optimal information disclosure), but it is analyzed in a more general setup (since the setup considered here covers cases with multiple agents, monetary instruments and any functional form of preferences as well as distributions of the state α). Compared to Kamenica and Gentzkow, it should be noted that they focus on when it is best for the Sender to disclose some information as opposed to none, which lies at the other extreme of the information disclosure question (as compared with the transparency question).

Note that when addressing the above question, I simply perform comparative statics varying the information structure of agent 1 (as all the papers just cited do). That is, I do not discuss the issue of how the information disclosure policy chosen by the designer would be interpreted by the agents so as to refine their estimate of α. Such a view seems appropriate to deal with organizations in which there is enough time to commit in advance (before the realization of α is known) to whatever disclosure policy sounds best. I also allow myself to choose the information structure the designer likes best. It may be argued that in some applications, some aspects of α cannot be hidden to the agent in which case the transparency question should be rephrased to allow only for admissible coarsening of agent 1's information. For the purpose of the main result below, α should be interpreted to stand for the characteristics of the problem that can potentially be hidden to agent 1.

It should be mentioned that the above abstract question can be related to the more concrete question as to whether it is good to disclose the individual contracts governing the various agents' incentives to their team partners. Indeed, not disclosing these contracts may be a way not to fully reveal α and also a way to hide some aspects of what shapes the team members' working incentives. Besides, to the extent that the agent's payoff does depend on α, the desirability of not disclosing α can be related to the possible advantage of using stochastic contracts rather than deterministic ones.

4.1 Full information benchmark

In the benchmark scenario, agents 1 and 2 know α (and in equilibrium they know each other's strategy). That is, given the instruments w, agents 1 and 2 play a Nash equilibrium of the complete information game defined by the payoff $u_i(a_1, a_2; w, \alpha)$ received by agent i for every action profile $a = (a_1, a_2)$ and the instrument(s) w.

In order to avoid technical complications, I will assume that u_i is a concave function of a that varies smoothly with w and α. Moreover, I will assume that whatever a_j,w, α, the function $a_i \to u_i(a_1, a_2; w, \alpha)$ is always maximized in a bounded subset of A_i (whatever w, α).

Such assumptions guarantee that 1) an interior pure strategy Nash equilibrium exists, and that 2) for almost all (w, α), Nash equilibria are locally unique and vary smoothly with (w,α) (see MasColell et al. (1995)).

I will denote by $a_i^{NE}(w,\alpha)$ one such equilibrium and I will assume it is the one describing the interaction in our team problem. Thus, in the benchmark scenario, in problem α, the designer sets the instruments $w = w(\alpha)$ (available to her) so as to maximize:

$$\pi(a_1^{NE}(w,\alpha), a_2^{NE}(w,\alpha); w, \alpha).$$

I will be interested in situations in which the solution obtained is different from the first-best solution the designer would choose if she could herself decide on a_1, a_2 as well as w. This is typically the case in moral hazard problems with one or two agents if transfers must be bounded and/or if agents are risk averse (unless agents' preferences are perfectly aligned with those of the designer and/or the designer can observe agents' actions and these are verifiable). Observe that in such cases, it is typically the case that $\bigtriangledown_{a_1}\pi(a_1^{NE}(w,\alpha), a_2^{NE}(w,\alpha); w, \alpha) \neq 0$. That is, even adjusting the instruments w optimally, the marginal effect of a_1 in the various directions need not be 0.

4.2 The coarse information case

To address the transparency question, I will consider situations in which agent 1 does not know whether $\alpha = \alpha_x$ or α_y while the designer and agent 2 do.[13] In this case, the relevant solution concept is the Bayes Nash equilibrium. The above concavity and smoothness assumptions guarantee that 1) there exists an interior pure strategy equilibrium and that 2) Bayes Nash equilibria (which are locally unique) inherit the smoothness properties of u_i for almost all w and α_x or α_y.

For each w_x, w_y, a Bayes Nash equilibrium is such that player 1 chooses action a_1^{CI} in both α_x, α_y and player 2 chooses actions $a_{2,x}^{CI}$ and $a_{2,y}^{CI}$ in α_x and α_y with

$$\begin{aligned}
a_{2,x}^{CI} &\in \arg\max_{a_2} u_2(a_1^{CI}, a_2; w_x, \alpha_x) \\
a_{2,y}^{CI} &\in \arg\max_{a_2} u_2(a_1^{CI}, a_2; w_y, \alpha_y) \\
a_1^{CI} &\in \arg\max_{a_1} p(\alpha_x) u_1(a_1, a_{2,x}^{CI}; w_x, \alpha_x) + p(\alpha_y) u_1(a_1, a_{2,y}^{CI}; w_y, \alpha_y)
\end{aligned}$$

[13] For the main result below, it is enough to consider information sets consisting of two states.

Letting $a^{CI}(w)$ denote the Bayes Nash equilibrium prevailing in the team problem, the best choice of instruments w is then obtained by maximizing

$$p(\alpha_x)\pi(a_1^{CI}(w), a_{2,x}^{CI}(w); w_x, \alpha_x) + p(\alpha_y)\pi(a_1^{CI}(w), a_{2,y}^{CI}(w); w_y, \alpha_y).$$

In the analysis, I will assume that if $a_1^{NE}(w_x, \alpha_x) = a_1^{NE}(w_y, \alpha_y)$ in the full information benchmark, then in the game in which agent 1 does not know whether α_x or α_y, the play is described by the complete information equilibrium strategy profile, as well. Clearly, such a strategy profile is a Bayes Nash equilibrium of the incomplete information game, and the assumption just made ensures that the comparison between the two informational scenarios is meaningful.[14]

A positive answer to question 1 is obtained when one can find α_x, α_y and $w^* = (w_x^*, w_y^*)$ such that

$$p(\alpha_x)\pi(a_1^{CI}(w^*), a_{2,x}^{CI}(w^*); w_x^*, \alpha_x) + p(\alpha_y)\pi(a_1^{CI}(w^*), a_{2,y}^{CI}(w^*); w_y^*, \alpha_y)$$

is strictly larger than

$$p(\alpha_x)\pi(a_1^{NE}(w_x, \alpha_x), a_2^{NE}(w_x, \alpha_x); w_x, \alpha_x) + p(\alpha_y)\pi(a_1^{NE}(w_y, \alpha_y), a_2^{NE}(w_y, \alpha_y); w_y, \alpha_y)$$

for all (w_x, w_y). In such a case, when agent 1 does not know $\alpha = \alpha_x$, α_y while agent 2 does and the available instruments w are set optimally, one can be sure that it delivers a strictly larger payoff to the organization than the one prevailing when both agents know α.

5 Main Result

The main result of this paper is that when the dimensionality s of α is sufficiently large, it is (generically) strictly desirable that at least one agent be coarsely informed of α. To present this result formally, I define the notion of genericity employed here. Let

[14] Alternatively, stronger conditions on u_i and u_j could be imposed that guarantee the uniqueness of the equilibrium.

$X = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{\omega_1+\omega_2} \times \mathbb{R}^s$ denote the domain of the profit functions π. Consider functions $\pi \in C^2(X)$. The set $\overline{\Pi}$ of $C^2(X)$ profit functions is endowed with a Whitney C^2 topology by letting a sequence $\pi_n \in \overline{\Pi}$ converge to π if and only if $\pi_n - \pi$ as well as the Jacobian of $\pi_n - \pi$ and the matrix of second derivative of $\pi_n - \pi$ converge uniformly to zero in the space of continuous functions over the relevant range of a_i, w, α (assumed to be bounded, see above). The definition of genericity is:

Definition. A set $\Pi \subseteq \overline{\Pi}$ is generic in $\overline{\Pi}$ if it contains a set that is open and dense in $\overline{\Pi}$.

The main result is:

Theorem 1 *Suppose the dimensionality of α is strictly bigger than the dimensionality of a_1, that is, $s > n_1$. Then for generic functions π, some non-full disclosure of α to agent 1 strictly enhances the designer's objective as compared with the full information benchmark.*

It should be mentioned that if one were to restrict attention to well-behaved profit functions π varying over countably many dimensions, say polynomials of any degree, then the conclusion of Theorem 1 would hold for a measure 1 set of π functions, thereby offering a measure-theoretic counterpart to this theorem. Besides, it should be highlighted that the conditions of Theorem 1 should be expected to be satisfied in most real life settings. Indeed, the parameter α characterizing the team problem should be thought of as containing at least information on the structure of the marginal cost incurred by each agent i along the various dimensions of his effort a_i (this has dimension no less than n_i), together with say information on the effect of the action profiles on the designer's objective, thereby making the dimension of the team problem s typically strictly larger than $n_1 + n_2 \geq \max(n_1, n_2)$. It should also be mentioned that whenever some aspects of α cannot be hidden to the agent (say because he knows it anyway - this could be relevant for the agent's own cost structure, for example), Theorem 1 still applies, as long as the dimension of the part of α that can be hidden to the agent exceeds the dimension of this agent's action space, which again sounds like the natural case in most applications of interest (since it seems possible to hide the cost structure of the other agent).[15,16]

5.1 The main arguments

Theorem 1 will first be established in the special case in which the principal has no instrument w, there is a single agent who has to choose a one-dimensional action, and the problem varies along two dimensions. It will then be explained how to extend the result to general multi-agent settings with arbitrary instruments for the designer and arbitrary dimensions n_1, n_2 whenever $s > n_1$.

Consider a setting with one agent whose action is $a \in \mathbb{R}$, and $\alpha \in \mathbb{R}^2$ parameterizes the agent's payoff function $u(a;\alpha)$. The complete information solution $a(\alpha)$ satisfies: $\frac{\partial}{\partial a}u(a;\alpha) = 0$.[17]

Consider α_0 in the interior of the α-space, and let

$$\overline{A}(\alpha_0) = \{\alpha \text{ such that } a(\alpha) = a(\alpha_0)\}.$$

For smooth u and generic α_0, this is a smooth (i.e. locally differentiable) manifold of dimension 1, which (locally around α_0) lies in the interior of the α-space.[18] Let $\alpha_1 \in \overline{A}(\alpha_0)$, $\alpha_1 \neq \alpha_0$ be in the interior of the α-space.

Starting from α_1, consider a direction δ in the α space in which $\alpha_1 + \varepsilon\delta$ is not in $\overline{A}(\alpha_0)$ for ε small enough and such that $\frac{\partial^2 u}{\partial a \partial \alpha^\delta}(a(\alpha_0);\alpha_1) \neq 0$. (Such a direction exists for generic

[15] For example, considering the moral hazard in team problem described in sections 2.2 and 3 (possibly in a multi-task version as in 2.1), the conclusion of Theorem 1 holds to the extent that the cost structure of agent j, $g_j(\cdot)$, together with the complementarity parameter β and the organizational reward R can be hidden to agent i and that the dimension of effort a_i is less than the dimension over which g_j, β and R can vary.

[16] Even in one-agent problems, playing on the possible monitoring technologies of the Principal would typically allow enough flexibility to get the conclusion of Theorem 1.

[17] Observe that what I am assuming here is only that (at least over a range of α) the solution $a(\alpha)$ varies smoothly with α and is locally pinned down by the first-order conditions (i.e., no other a in the neighborhood of $a(\alpha)$ satisfies the first-order condition). Thus, I am not following here the methodology of the first-order approach, as I am not maximizing the principal's objective assuming only that the first-order conditions for the agent are satisfied (see Hart and Holmström (1983) for some considerations on the first-order approach).

[18] The observation that $\overline{A}(\alpha_0)$ is a manifold of strictly positive dimension is key for the argument below and it is the key place where the dimensionality assumption $s > n_1$ is being used.

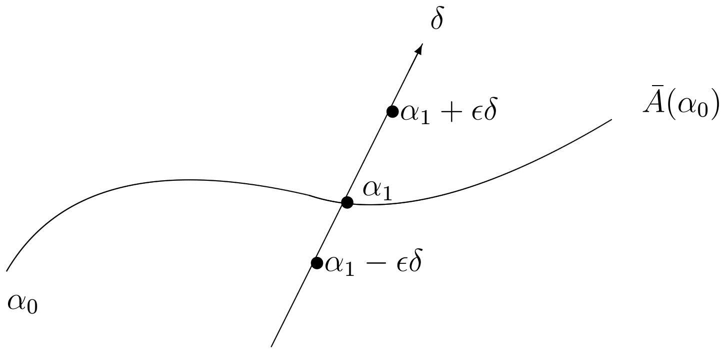

Figure 1:

u. For example, such a direction may be one in which the marginal effect of a on u is modified in proportion to a.)

Consider the problems $\alpha = \alpha_0$ and $\alpha_1 + \varepsilon\delta$ for ε either positive or negative but small (remember α_1 lies in the interior of the α-space). The idea is to compare the aggregate expected objective π when the agent knows whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$ and the expected objective when the agent ignores whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$. The various α considered here are illustrated in Figure 1.

Clearly for $\varepsilon = 0$, the two cases generate the same aggregate π by definition of $\overline{A}(\alpha_0)$.[19] But, for $\varepsilon \neq 0$, the two solutions will not in general lead to the same aggregate effect on π. I will now compute the first order effect in ε of this difference and show that it is generically different from 0, thereby allowing me to conclude that a coarse information of the above type either for $\varepsilon > 0$ and small or $\varepsilon < 0$ and small dominates the complete information case.

[19] In the general multi-agent extension, the selection hypothesis for Nash Bayes equilibria of games of incomplete information is being used as well.

Let $a_0 = a(\alpha_0)$ and $a_1(\varepsilon) = a(\alpha_1 + \varepsilon\delta)$. They satisfy

$$\begin{aligned}\frac{\partial}{\partial a}u(a_0;\alpha_0) &= 0 \\ \frac{\partial}{\partial a}u(a_1(\varepsilon);\alpha_1+\varepsilon\delta) &= 0\end{aligned} \tag{1}$$

Let $a^{CI}(\varepsilon)$ denote the action when the agent does not know whether $\alpha = \alpha_0$ or $\alpha_1+\varepsilon\delta$. It satisfies:

$$p(\alpha_0)\frac{\partial}{\partial a}u(a^{CI}(\varepsilon);\alpha_0) + p(\alpha_1+\varepsilon\delta)\frac{\partial}{\partial a}u(a^{CI}(\varepsilon);\alpha_1+\varepsilon\delta) = 0. \tag{2}$$

I wish to sign $\Delta(\varepsilon)$ defined as

$$p(\alpha_0)[\pi(a_0;\alpha_0) - \pi(a^{CI}(\varepsilon);\alpha_0)] + p(\alpha_1+\varepsilon\delta)[\pi(a_1(\varepsilon);\alpha_1+\varepsilon\delta) - \pi(a^{CI}(\varepsilon);\alpha_1+\varepsilon\delta)].$$

Clearly, if $\Delta(\varepsilon) < 0$, it is strictly better that the agent does not know whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$.

I now expand $\Delta(\varepsilon)$ at the first order in ε. Since $a^{CI}(0) = a_1(0) = a_0$, $\Delta(\varepsilon)$ writes at the first order:

$$p(\alpha_0)\frac{\partial\pi}{\partial a}(a_0;\alpha_0)[a_0 - a^{CI}(\varepsilon)] + p(\alpha_1)\frac{\partial\pi}{\partial a}(a_0;\alpha_1)[a_1(\varepsilon) - a^{CI}(\varepsilon)] + o(\varepsilon)$$

where $o(\varepsilon)$ denotes a function such that $\frac{o(\varepsilon)}{\varepsilon}$ goes to 0 as ε goes to 0.

Moreover from (1) and (2) (and using that $\frac{\partial^2 u}{\partial a^2} < 0$ is different from 0), we have that:

$$\begin{aligned}a_1(\varepsilon) - a_0 &= \frac{-\frac{\partial^2 u}{\partial a\partial\alpha^\delta}(\alpha_1)}{\frac{\partial^2 u}{\partial a^2}(\alpha_1)}\varepsilon + o(\varepsilon) \\ a^{CI}(\varepsilon) - a_0 &= \frac{-p(\alpha_1)\frac{\partial^2 u}{\partial a\partial\alpha^\delta}(\alpha_1)}{p(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_0) + p(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_1)}\varepsilon + o(\varepsilon)\end{aligned}$$

where $\partial h/\partial\alpha^\delta$ denotes the derivative of h in the direction α^δ and all functions are taken at $a = a_0$.

After multiplying $\Delta(\varepsilon)$ by $\frac{\partial^2 u}{\partial a^2}(\alpha_1)[p(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_0)+p(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_1)]$ and dividing by $p(\alpha_0)p(\alpha_1)$

(which are both strictly positive) we get that $\Delta(\varepsilon)$ has the same sign as

$$\left[\frac{\partial \pi}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \pi}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0)\right]\frac{\partial^2 u}{\partial a \partial \alpha^\delta}(\alpha_1)\varepsilon + o(\varepsilon)$$

Three cases may a priori occur.

1) $\left[\frac{\partial \pi}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \pi}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0)\right]\frac{\partial^2 u}{\partial a \partial \alpha^\delta}(\alpha_1) < 0$. Then taking $\varepsilon > 0$ and sufficiently small, we can infer from the above that not letting the agent know whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$ strictly dominates the complete information benchmark.

2) Likewise, if $\left[\frac{\partial \pi}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \pi}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0)\right]\frac{\partial^2 u}{\partial a \partial \alpha^\delta}(\alpha_1) > 0$, then taking $\varepsilon < 0$ and sufficiently small, not letting the agent know whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$ strictly dominates the complete information benchmark (remember than since α_1 is in the interior of the α-space, one can move in any direction from α_1).

3) The only case in which one cannot conclude is when

$$\left[\frac{\partial \pi}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \pi}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0)\right]\frac{\partial^2 u}{\partial a \partial \alpha^\delta}(\alpha_1) = 0$$

or

$$\frac{\partial \pi}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \pi}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0) = 0 \tag{3}$$

But, this condition is not satisfied for generic π functions.

To see this more formally, consider the family of π_λ functions

$$\pi_\lambda(a;\alpha) = \pi(a;\alpha) + \lambda a \parallel \alpha - \alpha_0 \parallel^2$$

where $\lambda \in \mathbb{R}$ and $\parallel \alpha - \alpha_0 \parallel$ denotes the euclidean distance between α and α_0. Obviously, if π satisfies (3), then for $\lambda \neq 0$, π_λ does not satisfy (3) -observe that changing π does not affect the expressions of $a_1(\varepsilon)$, $a^{CI}(\varepsilon)$- from which one can conclude that the set of π for which (3) does not hold is dense. Moreover, this set is also clearly open given the continuity of the mapping $\pi \to \frac{\partial \pi}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \pi}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0)$ according to the Whitney C^2 topology.[20]

[20] Clearly, if one were to consider polynomial functions π, then (3) would not hold for a measure 1 set of parameter values.

I now sketch how the argument extends to the general case considered in Theorem 1.

1) **Adding instruments w.**

Suppose the designer can now (optimally) choose instrument(s) w still assuming that there is a single agent. To fix ideas, take the above setting and assume that the designer can set $w \in \mathbb{R}$. For any α, there is an optimal w, say $w(\alpha)$. It is defined as

$$\begin{aligned} w(\alpha) &= \arg\max_{w} \pi(a(w), w; \alpha) \\ a(w) &= \arg\max_{a} u(a, w; \alpha) \end{aligned}$$

Thus, keeping α constant on has: $\frac{\partial u}{\partial a}(a, w; \alpha) = 0$, which after complete differentiation w.r.t w yields $\frac{\partial^2 u}{\partial a^2}\frac{da}{dw} + \frac{\partial^2 u}{\partial a \partial w} = 0$. The first-order condition on the designer's programme writes $\frac{\partial \pi}{\partial a}\frac{da}{dw} + \frac{\partial \pi}{\partial w} = 0$, which combined with the previous condition yields:

$$-\frac{\partial \pi}{\partial a}\frac{\frac{\partial^2 u}{\partial a \partial w}}{\frac{\partial^2 u}{\partial a^2}} + \frac{\partial \pi}{\partial w} = 0$$

Define $\overline{\pi}(a; \alpha) = \pi(a, w(\alpha); \alpha)$ and apply the argument developed above when there were no instruments assuming $\overline{\pi}$ is the designer's objective. Clearly, if not letting the agent know whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$ strictly dominates the complete information benchmark for this case, then in the case when the designer can choose w, it also strictly dominates (because the designer always has the option to set w to be $w(\alpha)$ in problem α).

It remains to show that generically it is not the case that (see (3) above)

$$\frac{\partial \overline{\pi}}{\partial a}(\alpha_0)\frac{\partial^2 u}{\partial a^2}(\alpha_1) - \frac{\partial \overline{\pi}}{\partial a}(\alpha_1)\frac{\partial^2 u}{\partial a^2}(\alpha_0) = 0 \tag{4}$$

To see this, consider the family of π_λ functions

$$\pi_\lambda(a; \alpha) = \pi(a; \alpha) + \lambda \parallel \alpha - \alpha_0 \parallel^2 (a - \frac{\partial^2 u/\partial a \partial w}{\partial^2 u/\partial a^2}(a_1, w(\alpha_1); \alpha_1)w)$$

where $\lambda \in \mathbb{R}$. For such a family, $w(\alpha)$ are the same at $\alpha = \alpha_0$ (resp. α_1) whatever λ so that $\frac{\partial \overline{\pi}_\lambda}{\partial a}(\alpha) = \frac{\partial \overline{\pi}}{\partial a} + \lambda \parallel \alpha - \alpha_0 \parallel^2$ for $\alpha = \alpha_0$ and α_1. Thus, if $\overline{\pi}$ satisfies (4), $\overline{\pi}_\lambda$ does not for any $\lambda \neq 0$, and one can conclude as before.

2) **Having more than one player.**

Roughly, this consists in extending the above differential arguments that were derived from one agent optimization conditions to a system of simultaneous optimization conditions as derived from the Nash equilibrium conditions.

Specifically, consider the case in which there is no instrument w. The FOC for NE (full information) write:

$$\begin{cases} \frac{\partial u_1}{\partial a_1}(a_1, a_2; \alpha) = 0 \\ \frac{\partial u_2}{\partial a_2}(a_1, a_2; \alpha) = 0 \end{cases}$$

which defines implicitly $a_1(\alpha)$ and $a_2(\alpha)$. Given that α has higher dimension than a_1 one can define (for generic u_1 and u_2) a manifold of dimension $s - n_1 \geq 1$ in the α space such that $a_1(\alpha) = a_1(\alpha_0)$, i.e. $\overline{A}(\alpha_0) = \{\alpha \text{ s.t. } a_1(\alpha) = a_1(\alpha_0)\}$.

Consider $\alpha_1 \in \overline{A}(\alpha_0)$ and a direction δ in the α space so that $\alpha_1 + \varepsilon\delta$ is known not to be in $\overline{A}(\alpha_0)$. If agent 1 does not know whether α_0 or $\alpha_1 + \varepsilon\delta$, NE actions a_1^C, $a_{2,0}^C$ and $a_{2,1}^C$ are given by:

$$\begin{cases} \frac{\partial u_2}{\partial a_2}(a_1^c(\varepsilon), a_{2,0}^c(\varepsilon); \alpha_0) = 0 \\ \frac{\partial u_2}{\partial a_2}(a_1^c(\varepsilon), a_{2,1}^c(\varepsilon); \alpha_1 + \varepsilon\delta) = 0 \\ p(\alpha_0)\frac{\partial u_1}{\partial a_1}(a_1^c(\varepsilon), a_{2,0}^c(\varepsilon); \alpha_0) + p(\alpha_1 + \varepsilon\delta)\frac{\partial u_1}{\partial a_1}(a_1^c(\varepsilon), a_{2,1}^c(\varepsilon); \alpha_1 + \varepsilon\delta) = 0 \end{cases}$$

And if there is full information, NE actions $a_{1,0}$, $a_{1,1}$, $a_{2,0}$ and $a_{2,1}$ are given by:

$$\begin{cases} \frac{\partial u_2}{\partial a_2}(a_{1,0}, a_{2,0}; \alpha_0) = 0 \\ \frac{\partial u_2}{\partial a_2}(a_{1,1}(\varepsilon), a_{2,1}(\varepsilon); \alpha_1 + \varepsilon\delta) = 0 \\ \frac{\partial u_1}{\partial a_1}(a_{1,0}, a_{2,0}; \alpha_0) = 0 \\ \frac{\partial u_1}{\partial a_1}(a_{1,1}(\varepsilon), a_{2,1}(\varepsilon); \alpha_1 + \varepsilon\delta) = 0 \end{cases}$$

I expand at order 1 in ε (the diff. of π in coarse vs full info)

$$\Delta(\varepsilon) = p(\alpha_0)[\pi(a_1^c(\varepsilon), a_{2,0}^c(\varepsilon);\alpha_0) - \pi(a_{1,0}, a_{2,0};\alpha_0)] + p(\alpha_1+\varepsilon\delta)[\pi(a_1^c(\varepsilon), a_{2,1}^c(\varepsilon);\alpha_1+\varepsilon\delta)] - \pi(a_{1,1}(\varepsilon), a_{2,1}(\varepsilon);\alpha_1+\varepsilon\delta)]$$

Similarly to the one agent case if $\Delta'(0) \neq 0$, then it implies that not letting agent 1 know whether α_0 or $\alpha_1 + \varepsilon\delta$ with $\varepsilon > 0$ or $\varepsilon < 0$ but small strictly improves over the full information benchmark and $\Delta'(0) = 0$ can be shown to be non-generic by considering perturbations of the form $\pi_\lambda(a_1, a_2;\alpha) = \pi(a_1, a_2;\alpha) + \lambda a_1 \parallel \alpha - \alpha_0 \parallel^2$.

5.2 Discussion

1) The logic of the argument used in subsection 2.1 does not quite follow the method of proof considered above when allowing for monetary instruments in Principal-agent interactions. Indeed, the method of proof consisted in freezing the monetary instruments at the full information optimal ones, and reason as in the case with no monetary instruments. By contrast, in Subsection 2.1, the monetary instruments were changed relative to the (optimal) full information ones, and this change together with the coarsening of information induced the strict improvement.[21]

Yet, the argument used in subsection 2.1 would work more generally. To illustrate this, consider the case in which the action space as well as the space of monetary instruments w are one-dimensional, while the space of α is of dimension 2. Consider $\alpha_1 \in \overline{A}(\alpha_0)$, $\alpha_1 \neq \alpha_0$ as in subsection 2.1 with w_0 and w_1 being the optimal (full information) values of w when $\alpha = \alpha_0$ or $\alpha = \alpha_1$, respectively.

The idea is to bundle α_0 and α_1 into one information set while allowing to move w away from the full information levels. Specifically, consider $w_1' = w_1 + \varepsilon$ where ε is assumed to be small (but may be either positive or negative) and keep w_0 unchanged. It turns out that for generic functions π, such a coarse information with a small ε (either positive or

[21] More precisely, in that problem one took $\alpha_y \in \overline{A}(\alpha_x)$ and one showed that by assigning α_x and α_y to the same information set and making some changes of the monetary instruments w one could generate strictly larger expected profit for the organization as compared with the full information benchmark (whereas the above argument required putting α_x and some α_y' in the neighborhood of α_y in the same information set and proceed with the full information optimal values of w.

negative) is strictly beneficial to the organization as compared with the full information benchmark.

In the coarse information case, the agent solves:

$$\max_a p(\alpha_0)u(a, w_0; \alpha_0) + p(\alpha_1)u(a, w_1 + \varepsilon; \alpha_1) \tag{5}$$

Let $a(\varepsilon)$ denote the action chosen by the agent and let

$$\widetilde{\pi}(\varepsilon) = p(\alpha_0)\pi(a(\varepsilon), w_0; \alpha_0) + p(\alpha_1)\pi(a(\varepsilon), w_1 + \varepsilon; \alpha_1)$$

denote the corresponding expected organization payoff. What we wish to show is that for generic values $\frac{d\widetilde{\pi}}{d\varepsilon}(\varepsilon = 0)$ is not 0, thereby ensuring that there exists a small ε (either positive or negative) so that $\widetilde{\pi}(\varepsilon) > \widetilde{\pi}(0)$ where $\widetilde{\pi}(0)$ is also the full information expected organizational payoff.

To show that $\frac{d\widetilde{\pi}}{d\varepsilon}(0) \neq 0$ (generically), write the first-order condition of (5) as:

$$p(\alpha_0)\frac{\partial u}{\partial a}(a, w_0; \alpha_0) + p(\alpha_1)\frac{\partial u}{\partial a}(a, w_1 + \varepsilon; \alpha_1) = 0.$$

Differentiation w.r.t ε (at $\varepsilon = 0$) yields:

$$[p(\alpha_0)\frac{\partial^2 u}{\partial a^2}(a, w_0; \alpha_0) + p(\alpha_1)\frac{\partial^2 u}{\partial a^2}(a, w_1; \alpha_1)]\frac{\partial a}{\partial \varepsilon} = -p(\alpha_1)\frac{\partial^2 u}{\partial a \partial w}(a, w_1; \alpha_1),$$

which in turn yields that $\frac{\partial \widetilde{\pi}}{\partial \varepsilon}$ is equal to:

$$[p(\alpha_0)\frac{\partial \pi}{\partial a}(a, w_0; \alpha_0)+p(\alpha_1)\frac{\partial \pi}{\partial a}(a, w_1; \alpha_1)]\frac{-p(\alpha_1)\frac{\partial^2 u}{\partial a \partial w}(a, w_1; \alpha_1)}{p(\alpha_0)\frac{\partial^2 u}{\partial a^2}(a, w_0; \alpha_0) + p(\alpha_1)\frac{\partial^2 u}{\partial a^2}(a, w_1; \alpha_1)}+p(\alpha_1)\frac{\partial \pi}{\partial w}(a, w_1; \alpha_1) \tag{6}$$

where the first-order conditions for the optimal determinations of w_0 and w_1 write

$$\begin{aligned}
\frac{\partial \pi}{\partial a}(a, w_0; \alpha_0)\frac{-\frac{\partial^2 u}{\partial a \partial w}(a, w_0; \alpha_0)}{\frac{\partial^2 u}{\partial a^2}(a, w_0; \alpha_0)} + \frac{\partial \pi}{\partial w}(a, w_0; \alpha_0) &= 0 \\
\frac{\partial \pi}{\partial a}(a, w_1; \alpha_1)\frac{-\frac{\partial^2 u}{\partial a \partial w}(a, w_1; \alpha_1)}{\frac{\partial^2 u}{\partial a^2}(a, w_1; \alpha_1)} + \frac{\partial \pi}{\partial w}(a, w_1; \alpha_1) &= 0
\end{aligned}$$

It is now not difficult to see that generically it cannot be that the expression in (6) is null.[22]

2) The above argument for Theorem 1 shows that one can gain by not letting agent 1 know whether $\alpha = \alpha_0$ or $\alpha_1(\alpha_0)$ ($\alpha_1(\alpha_0) = \alpha_1 + \varepsilon\delta$ in the above notation). By considering a positive mass neighborhood of $N(\alpha_0)$ and the corresponding $\alpha_1(\alpha)$ for $\alpha \in N(\alpha_0)$, one can in fact show that the gains of not letting agent 1 know whether $\alpha \in N(\alpha_0)$ or $\alpha_1(\alpha)$ are strictly positive in expected terms. Building on the example of subsection 2.1, it is not hard to verify that the gains of non-transparency can be made arbitrarily large (in relative terms), since in that example the complete information benchmark resulted in 0 aggregate profit whereas strictly positive profit could be achieved with coarse information.

3) Getting back to the trade-off (resulting from coarsening the information partition) between relaxing the incentive constraints (through aggregation) and constraining the strategy (through measurability constraints), Theorem 1 shows that one can always find an information partition such that the former effect dominates the latter. Yet, the argument used to prove this is not to show that the latter effect can be made of second order as compared with the former effect. In the construction, when agent 1 does not know whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$, both effects are of the same order. The result follows because, it is generically the case that for either $\varepsilon > 0$ or $\varepsilon < 0$ but small the former effect dominates the latter effect.

4) In the above analysis, I have implicitly ignored agents' participation constraints. Consider now imposing that agents should get at least their outside option payoff. Clearly, nothing changes if the participation constraints are not binding.[23] For example, in contexts with limited liability, agents typically receive a positive rent in moral hazard problems and the participation constraints are not binding. In the absence of limited liability constraints though, the designer would typically adjust the instruments w so that agents get their outside option payoff in pure moral hazard problems (see Holmström (1979-1982)

[22]Consider families of π such that $\frac{\partial\pi}{\partial a}(a, w_k; \alpha_k)/\frac{\partial\pi}{\partial w}(a, w_k; \alpha_k)$ remains constant for $k = 0, 1$ but $\frac{\partial\pi}{\partial a}(a, w_0; \alpha_0)/\frac{\partial\pi}{\partial a}(a, w_1; \alpha_1)$ varies. The full information conditions remain unchanged while the expression of $\frac{\partial\widetilde{\pi}}{\partial\varepsilon}$ varies (unless $\frac{\partial^2 u}{\partial a\partial w} = 0$, which for generic functions u would not hold).

[23]If the participation constraints are binding both at $\alpha = \alpha_0$ and $\alpha_1 + \varepsilon\delta$ in the main argument used to prove Theorem 1 when w is set at $w(\alpha)$ in problem α, one has to worry that the agent gets no less than his outside option payoff when the agent does not know whether $\alpha = \alpha_0$ or $\alpha_1 + \varepsilon\delta$, which may require increasing the burden to the designer.

or Holmström-Milgrom (1991) in the context of risk-averse agents without limited liability constraints). It should be noted however that if in addition to the moral hazard problem, agents were assumed to possess some private information then most "types" of agents would receive positive rent even in the absence of limited liability constraints. Theorem 1 could then be applied to such settings.

5) In the above framework, I have not allowed for mechanisms in which a mediator could make recommendations to agents as to which actions to choose. If such mechanisms are allowed, one can always implement the optimal mechanism by having agents be only informed of what to do (action a_i for agent i) (see Myerson, 1982). From that perspective, what Theorem 1 shows is the stronger property that when the dimension of α is larger than the dimension of agents' actions it *cannot* be optimal to let the agents know α.

6) In the context of Theorem 1, only the information structure of agent 1 was varied (as agent 2 was assumed to have complete information). If one further imposes that the information (about α) should be public among agents 1 and 2, then the same kind of non-transparency result as in Theorem 1 prevails, as long as the dimensionality of α is bigger than the sum of the dimensions of both agents' actions, i.e. $s > n_1 + n_2$. The idea is now to work with the manifold

$$B(\alpha_0) = \{\alpha \text{ s.t. } a_1(\alpha) = a_1(\alpha_0) \text{ and } a_2(\alpha) = a_2(\alpha_0)\}$$

which for generic α_0 has dimension $s - (n_1 + n_2)$.

6 Disclosure policy in low dimensional cases

In this Section, I consider settings in which the dimension of the state space is the same as the dimension of agents' action spaces, and I observe in representative classes of applications that full transparency is often the best disclosure policy in contrast to the message of Theorem 1 above.

6.1 Delegation with heterogeneous congruence

Consider the authority model of Aghion and Tirole (1997):

$$
\begin{aligned}
u_1(a_1, a_2; \alpha) &= a_2\beta b + (1 - a_2)a_1 b - g_1(a_1) \\
u_2(a_1, a_2; \alpha) &= a_2 B + (1 - a_2)a_1\gamma b - g_2(a_2)
\end{aligned}
$$

And assume that the sole source of heterogeneity is the parameter γ of congruence. I also let $b = B$, $\beta = 1$ and $g_1(a) = g_2(a) = \frac{a^2}{2}$.

I show that no matter what $\gamma_1 < \gamma_2 < ... < \gamma_n$ are, the principal (agent 2) is better off when the agent (agent 1) knows which γ is prevailing rather than when he does not know whether $\gamma = \gamma_1$, γ_2... or γ_n.

Routine calculations yield

$$
\begin{aligned}
a_1^{NE}(\gamma) &= \frac{B(1 - B)}{1 - \gamma B^2} \\
a_2^{NE}(\gamma) &= \frac{B(1 - \gamma B)}{1 - \gamma B^2}
\end{aligned}
$$

and when the agent does not know whether $\gamma = \gamma_1$.... or γ_n (while the principal does):

$$
\begin{aligned}
a_1^{CI} &= \frac{B(1 - B)}{1 - E(\gamma)B^2} \\
a_2^{CI}(\gamma_i) &= B[1 - \gamma_i \frac{B(1 - B)}{1 - E(\gamma)B^2}]
\end{aligned}
$$

where $E(\gamma)$ is the expected value of γ.

Given the convexity of $\gamma \rightarrow \frac{B(1-B)}{1-\gamma B^2}$, it is readily verified that $E(a_1^{NE}(\gamma)) > a_1^{NE}(E(\gamma)) = a_1^{CI}$. Furthermore, as common sense suggests, agent 2's effort decreases with the degree of congruence in the coarse information case $a_2^{CI}(\gamma_1) > a_2^{CI}(\gamma_2).... > a_2^{CI}(\gamma_n)$, and agent 1's effort increases with the degree of congruence γ in the full information case $a_2^{NE}(\gamma_1) < a_2^{NE}(\gamma_2).... < a_2^{NE}(\gamma_n)$.

The difference of agent 2's expected payoff in the coarse information case and the complete information case writes:

$$\sum_i p(\gamma_i) \max_{a_2}[Ba_2 + B(1-a_2)\gamma_i a_1^{NE}(\gamma_i) - g_2(a_2)] -$$
$$\sum_i p(\gamma_i) \max_{a_2}[Ba_2 + B(1-a_2)\gamma_i a_1^{CI} - g_2(a_2)]$$

It is no smaller than $\sum_i p(\gamma_i)B(1 - a_2^{CI}(\gamma_i))\gamma_i(a_1^{NE}(\gamma_i) - a_1^{CI})$ (because in the max appearing in the first summation on can always pick $a_2 = a_2^{CI}(\gamma_i)$ when $\gamma = \gamma_i$, i.e. the argument for the corresponding max in the second summation). This is itself no smaller than

$$\sum_i p(\gamma_i)B(1 - a_2^{CI}(\gamma_i))\gamma_i(a_1^{NE}(\gamma_i) - E(a_1^{NE}(\gamma))$$

which is strictly positive given that $\gamma \to (1-a_2^{CI}(\gamma))\gamma$ and $\gamma \to a_1^{NE}(\gamma)$ are both increasing with γ.

To summarize,

Proposition 1 *In the optimal delegation problem with quadratic cost of effort, when the sole heterogeneity is on the congruence parameter γ, full disclosure of γ is always the best policy for the principal.*

The intuition for this result is as follows. Not letting the agent know γ leads him to pick his effort level as a best-response to a mixed distribution of Principal's effort. This in turn leads the agent to make less effort than in the complete information case when the congruence parameter γ is bigger and more effort when it is smaller. But, the Principal would prefer the bias to be the other way round given the implication of the congruence parameter, thereby explaining why full information disclosure is preferable in this case.

Several remarks are in order regarding this proposition. First, even though Proposition 1 was established for the case in which the organizational objective coincides with agent 2's payoff, it should be clear that the same conclusion continues to hold for organizational objective functions that would lie in a neighborhood of agent 2's objective function, thus the conclusion holds generically. Second, the result of Proposition 1 is not in contradiction with Theorem 1 above because the setup analyzed here is one in which the dimensionality of α is the same as the dimensionality of the effort of the agent (so that there is no manifold of strictly positive dimension in the α space in which at the Nash equilibrium, the agent performs the same effort level).[24]

6.2 Moral hazard in team with heterogeneous complementarity

Consider the following moral hazard in team problem in which agent $i = 1, 2$'s payoff is

$$u_i(a_1, a_2; \beta) = (a_1 + a_2 + \beta a_1 a_2)\, w - \frac{(a_i)^2}{2}$$

and the corresponding profit is

$$\pi(a_1, a_2; \beta) = (a_1 + a_2 + \beta a_1 a_2)\,(R - 2w).$$

Assume that the sole degree of heterogeneity is the complementarity parameter $\beta \in \left[\underline{\beta}, \overline{\beta}\right]$. I simplify the analysis by assuming that the bonus w is not an instrument of the designer, and that it is set independently of β and satisfies $w < R/2$.[25]

I consider symmetric information disclosure policies for the two agents. Specifically, let $\beta^1 < \beta^2 .... < \beta^n$ and let p^k denote the probability of β^k. Consider both the case of complete information and the case of coarse information in which no agent $i = 1$ or 2 knows whether β is $\beta^1, ...$or β^n.

Proposition 2 *The coarse information disclosure policy always generates strictly less expected profit to the designer than the complete information disclosure policy.*

The intuition for this result is as follows. Not letting the agents know which β prevails leads agents to make comparatively more effort when the complementarity parameter is low and less effort when it is large. This is bad for the overall profit because the marginal effect of effort is larger when the complementarity parameter is larger, thereby explaining why the complete information disclosure policy dominates in this case. The detailed proof of Proposition 2 appears in Appendix. Observe again that this result is

[24] Reproducing the argument for Theorem 1 with $\alpha_1 = \alpha_0$ would yield that $\Delta(\varepsilon)$ is of the same order as ε^2, and thus one would not be able to conclude from the argument given there.

[25] Such an assumption would fit if we have in mind that the bonus w is negotiated after a success is being obtained and the two agents have the same bargaining power (independent of β).

not in contradiction with the insight of our main result (Theorem 1) given that here the dimensionality of the problem is equal to the dimensionality of the effort level.

6.3 Sender-Receiver interaction with uniform bias

Consider the Sender-Receiver (Crawford-Sobel, 1982) model with quadratic utility functions. That is, the state ω varies in $[0,1]$ according to some smooth distribution with density $f(.)$. The receiver chooses an action a in $[0,1]$. The receiver and sender's payoffs depend on a and ω according to

$$\begin{aligned} u_R(a;\omega) &= -(a-\omega)^2 \\ u_S(a;\omega) &= -(a-\omega-b)^2 \end{aligned}$$

The quadratic payoff of the receiver implies that the receiver chooses an action equal to the expected value of ω (given his information). The Sender on the other hand would like the action to be biased toward higher (resp lower) values whenever $b>0$ (resp $b<0$).

Given the strict concavity of the function $x \to -(x-b)^2$, Jensen's inequality implies that whatever the non-degenerate distribution of ω,

$$-E[E(w)-w-b]^2 < -b^2 \tag{7}$$

(since $E[E(w)-w]=0$). Given that the left-handside is what the Sender would get in expectation in a coarse disclosure case and the right-handside is what she gets in the complete information case, (7) implies that:

Proposition 3 *In the Sender-Receiver interaction with quadratic utility functions, full transparency is always best.*

Again, the above result does not contradict our main result, since the setup considered in the above proposition is one in which the dimension of the state ω coincides with the dimension of the Receiver's action. As already mentioned in Introduction, Rayo and Segal (2010) consider a case in which the information held by the Sender is two-dimensional and they observe in their setup that full disclosure is never best in accordance with the main result of this paper.

7 Conclusion

In this paper, I have shown that some form of non-transparency is desirable, as long as the dimensionality of what can be hidden exceeds the dimensionality of agents' action spaces. It is tempting to relate the theoretical result obtained in this paper to the widespread practice in organizations not to inform workers of other workers' wages and bonus schemes. More work though should be devoted to analyze in more details the pros and the cons of such practices. More generally, it would be desirable to make progress on which characteristics of the problem α should be kept secret and which ones should be made publicly available to agents.

Appendix (Proof of Proposition 2)

Routine calculations yield that in the full information case agents choose $a^{NE}(\beta) = \frac{w}{1-\beta w}$ when the complementarity parameter is β. In the coarse information case, agents choose:

$$a^{CI} = \frac{w}{1 - E(\beta)w}$$

where $E(\beta)$ denotes the expected value of β. Given the convexity of $\beta \to \frac{w}{1-\beta w}$, it follows by Jensen's inequality that

$$a^{CI} < E(a^{NE}(\beta)). \tag{8}$$

The difference of expected profit in the complete disclosure case and in the coarse information case writes:

$$\begin{aligned} \Delta/(R-2w) &= \sum\nolimits_i p^i(2a^{NE}(\beta^i) + \beta^i(a^{NE}(\beta^i))^2 - 2a^{CI} - \beta^i(a^{CI})^2 \\ &= 2\sum\nolimits_i p^i(a^{NE}(\beta^i) - a^{CI}) + \sum\nolimits_i p^i\beta^i((a^{NE}(\beta^i))^2 - (a^{CI})^2) \end{aligned}$$

We have that $\sum_i p^i(a^{NE}(\beta^i) - a^{CI}) > 0$ by (8).

Moreover, let $i^* = \arg\min_i \left\{i \text{ such that } a^{NE}(\beta^i) \geq a^{CI}\right\}$. Given the monotonicity of $i \to a^{NE}(\beta^i)$, we have that for $i \geq i^*$, $a^{NE}(\beta^i) \geq a^{CI}$ and for $i < i^*$, $a^{NE}(\beta^i) \leq a^{CI}$. Writing $(a^{NE}(\beta^i))^2 - (a^{CI})^2$ as $(a^{NE}(\beta^i) + a^{CI}))(a^{NE}(\beta^i) - a^{CI})$, making use of the monotonicity of $i \to \beta^i \left(a^{NE}(\beta^i) + a^{CI}\right)$, and of the change of sign of $a^{NE}(\beta^i) - a^{CI}$ at $i = i^*$, we get that for all i

$$\beta^i \left((a^{NE}(\beta^i))^2 - (a^{CF}(\beta^i))^2\right) \geq \beta^{i^*}(a^{NE}(\beta^{i^*}) + a^{CF}(\beta^{i^*}))(a^{NE}(\beta^i) - a^{CF}(\beta^i))$$

In turn, this implies that

$$\begin{aligned} &\sum\nolimits_i p^i\beta^i((a^{NE}(\beta^i))^2 - (a^{CI})^2) \\ \geq\ &\beta^{i^*}(a^{NE}(\beta^{i^*}) + a^{CI})\sum\nolimits_i p^i(a^{NE}(\beta^i) - a^{CI}) \end{aligned}$$

which again is strictly positive by (8). **Q. E. D.**

References

[1] Abreu, D, P. Milgrom and D. Pearce (1991): 'Information and Timing in Repeated Partnerships,' *Econometrica* **59**, 1713-1733.

[2] Aghion, P. and J. Tirole (1997): 'Formal and Real Authority in Organizations,' *Journal of Political Economy*, **105**, 1-29.

[3] Aoyagi, M. (2010): 'Information Feedback in a Dynamic Tournament,' *Games and Economic Behavior*, **70**, 242-260.

[4] Baker, G. (2010): 'The Use of Performance Measures in Incentive Contracting,' *American Economic Review* **90**, 415-420.

[5] Bassan, B., O. Gossner, M. Scarsini and S. Zamir (2003): 'Positive value of information in games,' *International Journal of Game Theory* **32**, 17-31.

[6] Battaglini, M. (2006): 'Joint Production in Teams,' *Journal of Economic Theory*, **130**, 138-167.

[7] Crawford, V. and J. Sobel (1982): 'Strategic Information Transmission,' *Econometrica*, **50**, 1431-1451.

[8] Crémer, J. (1995): 'Arm's Length Relationships' The *Quarterly Journal of Economics*, **110**, 275-295.

[9] Dewatripont, M., I. Jewitt and J. Tirole (1999): 'The Economics of Career Concerns Part I: Comparing Information Structures,' *Review of Economic Studies* **66**, Special Issue, 199-217.

[10] Dubey, P. and J. Geanakoplos (2010): 'Grading Exams: 100,99,98,... or A,B,C?' *Games and Economic Behavior* **69**, 72-94.

[11] Ederer, F., R. Holden and M. Meyer (2008): 'Gaming and Strategic Ambiguity in Incentive Provision', mimeo.

[12] Gavazza A. and A. Lizzeri (2009): 'Transparency and Economic Policy,' *Review of Economic Studies* **76**, 1023-1048.

[13] Grossman, S. and O. Hart (1986): 'The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,' *Journal of Political Economy* **94**, 691-719.

[14] Hart, O and B. Holmström (1983): 'An Analysis of the Principal-Agent Problem,' *Econometrica* **51**, 7-45.

[15] Hirshleifer, J. (1971): 'The Private and Social Value of Information and the Reward to Inventive Activity,' *American Economic Review* **61**, 561-574.

[16] Holmström, B (1979): 'Moral Hazard and Observability' The *Bell Journal of Economics*, **10**, 74-91.

[17] Holmström, B. (1982): 'Moral Hazard in Teams' The *Bell Journal of Economics*, **13**, 324-340

[18] Holmström, B. and P. Milgrom (1991): 'Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design' *Journal of Law, Economics, and Organization* **7**, 24-52.

[19] Kamenica, E. and M. Gentzkow (2011): 'Bayesian persuasion,' *American Economic Review*, **101**, 2590-2615.

[20] Lizzeri, A., M. Meyer and N. Persico (2002): 'The Incentive Effects of Interim Performance Evaluations,' Caress Working Paper.

[21] Mas-Colell, A., M. D. Whinston, and J. R. Green (1995): *Microeconomic Theory* Oxford University Press.

[22] Milgrom, P. and R. Weber (1982): 'A Theory of Auctions and Competitive Bidding,' *Econometrica*, **50**, 1089-1122.

[23] Myerson, R. (1982): 'Optimal Coordination Mechanisms in Generalized Principal-Agent Problems,' *Journal of Mathematical Economics*, **10**, 6781.

[24] Ottaviani, M. and A. Prat (2001): 'The value of public information in monopoly,' *Econometrica*, **69**, 1673-1683.

[25] Prat, A. (2005): 'The Wrong Kind of Transparency,' The *American Economic Review*, **95**, 862-877

[26] Rahman, D. (2010): 'But Who Will Monitor the Monitor,' mimeo University of Minnesota.

[27] Rayo, L. and I. Segal (2010): 'Optimal information disclosure,' *Journal of Political Economy*, **118**, 949-987.

[28] Winter, E. (2010): 'Transparency among Peers and Incentives,' forthcoming *Rand Journal of Economics*.