On Transparency in Organizations*

Philippe Jehiel†

8th January 2014

Abstract

When is it best for the Principal to commit not to disclosing every piece of information she may have to the agent in moral hazard interactions. I show that whenever there exist at least two distinct states in which the agent would choose the same action in the complete information benchmark, then full transparency can be improved upon. It is suggested that such situations are common, and the optimal way of remaining silent is characterized in the context of a simple moral hazard interaction in which the state is parameterized by how difficult the task is and how good the monitoring technology is.

1 Introduction

A central question of economics is about how to provide the best incentives to agents in an attempt to improve the working of organizations. Most of the literature in contract theory concerned with this question has focused on how to use monetary instruments so as to best align the objectives of the agents with those of the organization at a minimal

*Thanks to seminar participants at Cesifo IO workshop, Paris 1, the workshop on multidimensional mechanism design in Bonn, Warwick workshop, StonyBrook game theory conference, Queen Mary University workshop, the conference on gaming in Bonn, Sciences Po, and in particular Ariane Lambert, Laurent Lamy, David Martimort, Paul Milgrom, Andy Newman, Larry Samuelson and Wataru Tamura for insightful comments. I also thank the European Research Council for financial support, and finally the editor (Marco Ottaviani) and the referees for helpful comments.

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

cost. But, an equally important design question concerns the distribution of information in organizations insofar as information affects the perceived consequences of actions and thus the incentives of agents.

Consider a Principal-agent relationship in which the agent must exert some effort so as to increase the chances of success of the task. The Principal does not observe effort directly, and she is informed at the time the task is to be completed about how difficult it is to achieve a successful outcome and how noisy it is to measure success in a reasonable time scale where the agent is assumed to enjoy some exogenously set bonus in case a positive signal about success is received ex post by the Principal.

A possible disclosure policy that I refer to as full transparency requires that the agent be informed prior to making his effort decision about the monitoring technology in the hands of the Principal and how difficult the task is. Alternative less transparent disclosure policies would require that some aspects of the information held by the Principal be kept unknown to the agent, as least until the completion of the task. The question studied in this paper is whether full transparency should be expected to be the disclosure policy that is most preferred by the Principal in such situations.

At some naive level, it would seem that irrespective of what the Principal observes about the difficulty of the task, it is a good idea for the Principal when the monitoring technology is poor not to let the agent know about it, given that a poor monitoring technology translates in low incentives to exert effort for the agent. But, if the agent makes rational inferences as to what being uninformed implies, as this paper assumes, it is unclear such forms of non-transparency would be effective, given that being uninformed about the quality of the monitoring technology would only average the perceived quality of the incentive scheme (and thus transfers effort levels from some states in which the monitoring technology is good to other ones in which it is less good with unclear overall effects).

The main result of this paper will imply that in situations like the moral hazard one just described in which the information held by the Principal varies over more dimensions than the action of the agent, full transparency is not the optimal disclosure policy no matter what the exact functional forms describing the preferences, costs, physical and monitoring technologies are. It should be highlighted that the conclusion that full transparency is not optimal crucially relies on the possibility for the Principal to hide aspects that concern both the quality of the monitoring technology and the difficulty of the task, and it would not necessarily hold if the information held by the Principal were restricted to the quality of the monitoring technology (in contrast to the naive approach suggested above).[1] In other words, full transparency is not optimal in this moral hazard situation but for reasons that are not correctly captured by the naive intuition suggested above. I will later on elaborate on the intuition for this conclusion.

Specifically, I consider the transparency question in the context of the following general abstract model. An agent is engaged in a moral hazard interaction parameterized by a state θ. He must choose an action a. Both the state θ and the action a can be varied locally. In state θ, action a results in an expected payoff $u(a;\theta)$ to the agent and an expected payoff $\pi(a;\theta)$ to the Principal (or the organization). Full transparency would require that the agent be fully informed of the state θ before making his choice of action a. The question I am interested in is whether other less transparent disclosure policies could not be preferable for the organization. Or to put it differently, whether in expectation the Principal cannot achieve a better outcome (from the viewpoint of the organization) by having the agent be incompletely rather than completely informed of the state θ.

In addressing the above transparency question, I have in mind that the distribution of information about θ provided to the agent is decided at an ex ante stage before the realization of θ is known to the Principal. Thus, the choice of disclosure policy does not signal anything about which θ is being observed by the Principal in the current interaction of interest.[2] I believe such a scenario fits in well with the application to organizations in which due to a constant flow of new pieces of information the design of how information is distributed is better thought of from an ex ante perspective.

In principle, one could consider arbitrary disclosure policies when addressing the above transparency question, but in some applications, some choices of disclosure policy would

[1] In particular, full transparency would be optimal in such a case with a fully rational agent whenever the objective of the Principal is a sufficiently convex function of the effort of the agent.

[2] If the disclosure policy were chosen after the Principal observes the state θ, the work of Crawford and Sobel (1982) on strategic information transmission suggests that not every information held by the Principal would be transmitted to the agent, as soon as there is some conflict of interest between the Principal and the agent (the Sender and the Receiver in their words). As I observe later, when addressing the transparency question from an ex ante perspective as in this paper, full transparency is optimal in the leading quadratic example considered by Crawford and Sobel.

seem harder to implement (because they may require information devices that sound less natural). Having this constraint in mind, I consider the case in which the Principal has only one option: Either disclose fully the state θ to the agent or remain silent and not say anything about the state.

The main result of this paper can be stated as follows. For generic organizational objectives π, the Principal can do strictly better than full transparency by remaining silent over a well chosen subset of states, as long as there exist at least two distinct states θ_1 and θ_2 such that the action of the agent would be the same at these two states in the fully transparent benchmark.

Before I elaborate on the result, note that in the Principal-agent illustration suggested above, the same lower effort would be exerted by the agent in the full transparency benchmark either in the case of a less good monitoring technology of the Principal (making it harder to incentivize the agent) or in the case of a less strong causality between effort and success, thereby ensuring that my general non-transparency result applies to this situation.

More generally, there will typically be several distinct states that would induce the same choice of action in the fully transparent benchmark, as long as the dimension of the state space exceeds the dimension of the action space of the agent. For such situations, my main result says that full transparency is not optimal, and that full transparency can be improved upon using a fairly simple disclosure device that only requires that the Principal remains silent about the realization of the state for a well chosen subset of states.

Importantly, the same non-transparency result extends to situations in which in addition to the disclosure policy, the Principal can use monetary instruments to incentivize the agent (as most of the contract theory assumes) and to situations with more than one agent, covering applications such as moral hazard in teams. In all cases, full transparency can be improved upon, as long as there exist at least two distinct states that would induce the same choice of actions of the agents in the fully transparent benchmark.

Let me now provide some intuition as to why full transparency can be improved upon when the same action a would be chosen at two distinct states θ_1 and θ_2. Assume, for simplicity, that the action of the agent varies over one dimension, and consider a state θ_2^+ in the vicinity of θ_2 such that in the complete information benchmark the action at θ_2^+

is slightly larger than that at θ_2 and θ_1. Not letting the agent know whether the state is θ_1 or θ_2^+ is good for the Principal whenever either the principal cares more about the agent exerting more effort at θ_1 than at θ_2 (and the utility function of the agent is equally concave at θ_1 and θ_2^+) or the utility function of the agent is sufficiently more concave in effort at θ_2^+ than at θ_1 so that the effort in the nontransparent case is close to that at θ_2^+. I provide later a sense of how the relative sensitivity of the principal's objective to the agent's effort and the relative concavity in effort of the agent's utility should be aggregated for the Principal to prefer that the agent be uninformed as to whether the state is θ_1 or θ_2^+. Whenever this condition does not hold, I show that for generic specifications of the principal's objective, the Principal finds it strictly beneficial not to let the agent know whether θ is θ_1 or θ_2^- where θ_2^- is defined so that θ_2 is in the middle between θ_2^- and θ_2^+.[3] Thus, for generic organizational objectives, the Principal is strictly better off either when the agent does not whether $\theta = \theta_1$ or θ_2^+ or when he does not know whether $\theta = \theta_1$ or θ_2^- as compared with the full transparency benchmark. Completing the proof that the Principal can gain in expected terms over the full transparency benchmark requires that the set of states over which the Principal remains silent be of positive measure, and such a result can be obtained using simple continuity arguments.

The above result does not say how much can be gained over full transparency. I illustrate through the above moral hazard example that the gains over full transparency can be arbitrarily large in relative terms. In the final part of the paper, I also characterize when the Principal should optimally remain silent about the state in the context of the same leading example. In short, I show that the Principal should optimally remain silent when, on the one hand, the task is easy and the monitoring technology is poor and, on the other hand, when the task is difficult and the monitoring technology is good, thereby

[3] This results from the observation that the comparative advantage of not letting the agent know whether $\theta = \theta_1$ or θ_2^+ is of order $(\theta_2^+ - \theta_2)$ for generic functions π so that not letting the agent know whether $\theta = \theta_1$ or θ_2^- has an effect of of the opposite sign as not letting not letting the agent know whether $\theta = \theta_1$ or θ_2^+.

Observe how the argument relies on the assumption that for two distinct states, the same action would occur. Replacing state θ_2 by state θ_1 would result in a comparative advantage of not letting the agent know whether $\theta = \theta_1$ or θ_1^+ of the order of $(\theta_1^+ - \theta_1)^2$ (and not of the order of $(\theta_2^+ - \theta_2)$ when $\theta_2 \neq \theta_1$). It would thus be of the same sign as the comparative advantage of not letting the agent know whether $\theta = \theta_1$ or θ_1^-. As a result, it would not be possible to conclude that full transparency can be improved upon.

giving a precise sense of how the disclosure about the two dimensions of the state should be optimally combined in this application.

The transparency question addressed in this paper is closely related to the optimal information disclosure question considered in Rayo and Segal (2010) or Tamura (2012) and to the Bayesian persuasion question considered in Kamenica and Gentzkow (2011). These authors (mostly) consider Sender-Receiver interactions without monetary instruments 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). These papers ask in specific contexts: Which disclosure policy the Sender should commit to so as to maximize her expected payoff? It should be noted that these papers do not consider the possibility that the only option for the Sender is to remain silent or disclose fully the state.

Rayo and Segal observe in their setup that transparency is not best when there are no monetary instruments. Their result can be viewed as offering an illustration of the main non-transparency result of this paper. In Rayo and Segal, the action of the Receiver can be described as the probability of accepting the project offered by the Sender (even though in their model this action is derived from the random realization of the Receiver's outside option), and the state can be described as the profile of payoffs the Sender and the Receiver would receive in case the project would be adopted. Clearly, for a given Receiver's payoff of implementing the project, the probability of acceptance of the Receiver is the same irrespective of how valuable the project is to the Sender. Thus the key condition of the main Proposition of this paper is satisfied, thereby explaining why full transparency can be improved upon in Rayo and Segal' setup. In a later part of their paper, Rayo and Segal observe that when the Sender can use monetary transfers, full transparency becomes optimal. Yet, the transparency result they obtain when side-payments are allowed is not general and hinges on the specific functional forms that they consider in their paper, as the non-transparency result derived in Section 5 shows. While Rayo and Segal have offered an insightful exploration of the optimal disclosure policy in an interesting example, I believe the general treatment offered in this paper provides a better account of why full transparency should be expected to be suboptimal in general moral hazard interactions with or without monetary instruments when the dimension of the state space exceeds the dimension of the action space.[4]

Kamenica and Gentzkow (2011) mostly consider situations in which the state and the action take values over discrete realizations, and they note that the optimal disclosure policy typically involves a well adjusted noisiness in the precision of the transmission so as to optimally let the agent be indifferent over several actions in some instances. While Kamenica and Gentzkow (2011) provide a general way to cope with such investigations, I note that my setup makes an essential use of the possibility that the action and the state be locally varied, which plays no role in Kamenica and Gentzkow. I also note that their focus on Bayesian persuasion or on when it is best for the Sender to disclose some information (as opposed to none) to the Receiver lies at the other extreme of the non-transparency question considered in this paper.

The question addressed in this paper is also tangentially related to other strands of literature. For example, some papers have shown in adverse selection environments such as auctions or monopoly that it is best for the designer to transmit as much information as possible to the agent(s) whenever the information of the designer is affiliated with the information of the agent(s).[5] The main non-transparency result of this paper suggests in moral hazard environments that when the information of the designer has more dimensions than the action of the agent, there is no way (natural or not) to make full transparency optimal.

The rest of the paper is structured as follows. A general Principal-agent framework is presented in Section 2. Preliminary considerations together with a preliminary investigation of a leading moral hazard example are gathered in Section 3. Section 4 states, and discusses the main non-transparency result. Section 5 considers the extensions to multi-agent settings and to the case in which side-payments are allowed. Section 6 considers the optimal way to stay silent in the context of the leading example introduced in Section 3. Section 7 concludes.

[4] In Section 5, I also discuss why the insight about the optimality of full disclosure in exchange for well adjusted side-payments obtained by Eso and Szentes in an auction setup does not apply in general here (essentially because the full information disclosure would not in general result in the maximization of total welfare in typical moral hazard situations unlike in the auction setup).

[5] The relevant reference for auctions is Milgrom and Weber (1982), and for monopoly it is Ottaviani and Prat (2001).

2 A Principal-Agent framework

I consider a family of moral hazard problems with one agent parameterized by a state variable $\theta \in \mathbb{R}^m$. The state θ is assumed to be distributed according to a smooth (i.e., continuously differentiable) density $p(\theta)$ that is strictly positive on some open bounded subset of $\mathbb{R}^m$. In every state θ, the agent chooses an action a in A, an open subset of $\mathbb{R}^n$.

In state θ, the expected payoff to the agent is $u(a;\theta)$ when the agent chooses a. The corresponding expected payoff to the Principal is denoted by $\pi(a;\theta)$. I assume that u is a concave function of a that varies smoothly with θ. Moreover, I assume that whatever θ, the function $a \to u(a;\theta)$ is always maximized in a bounded subset of A.[6] The function π is also assumed to be smooth.

The general theme I wish to explore is whether it is in the interest of the Principal that the agent be informed of θ whatever its realization before he makes his choice of action a. Specifically, I ask:

Question 1. When is it beneficial for the designer that the agent be incompletely informed of θ? Or to put it differently, when is some form of non-transparency desirable?

When some form of non-transparency is desirable, it would be of interest to know more about the best disclosure policy for the Principal. In some parts of the paper (when putting more structure on the problem), I address the following question:

Question 2. Assuming the Principal can ex ante commit for each realization of θ to either telling what the realization of the state θ is or remaining silent about this, what is the best strategy for the Principal?

When addressing the above questions, observe that I am assuming that the Principal can commit in advance (before knowing the realization of θ) to the chosen disclosure policy. This implies that the choice of information disclosure policy cannot help the agent in his estimate of θ, since that choice is made before the Principal knows θ. This is the same ex ante view as the one adopted in Rayo-Segal (2010), Kamenica-Gentzkow (2011), and it seems appropriate to deal with organizations in which there is enough time to

[6] Such an assumption will typically guarantee that one can work with first-order conditions to deal with the maximization problem of the agent.

commit in advance (before the realization of θ is known) to whatever disclosure policy sounds best.

While question 2 does not deal with the most general form of communication strategy, I believe that the option of remaining silent is a common (and widespread) communication device. It is also fairly easy to implement. Of course, the premise of the analysis is that the agent would make the correct inference as to how the state θ is distributed when the Principal remains silent.

3 Preliminaries

3.1 Preference alignments

Clearly, if the Principal and the agent have the same preferences, $\pi = u$, then full transparency is best because in a decision problem one can only benefit from having more information. In such a case, the best disclosure policy is for the Principal to tell the agent the realization θ.

Symmetrically, if the principal and the agent have completely opposed preferences in the sense that there exists a constant c such that for all a and θ, $u(a;\theta) + \pi(a;\theta) = c$, then remaining completely silent as to what the realization of θ is is the best disclosure policy for the Principal, given that the more informed the agent, the better for him and hence the worse for the Principal in this zero-sum context.

From these two basic observations, it would seem that the preference alignment between the Principal and the agent is the main driving force behind the desirability of transparency. Yet, consider the quadratic example of Crawford and Sobel (1982) in which preferences can be represented as:

$$\begin{aligned} u(a;\theta) &= -(a-\theta-b)^2 \\ \pi(a;\theta) &= -(a-\theta)^2 \end{aligned}$$

The principal would like the action a of the agent to be as close as possible to the state θ whereas the agent would like this action to be as close as possible to $\theta + b$ where b can be interpreted as measuring the size of the bias of the agent or the degree of misalignment of the preferences of the agent and the Principal.

Interestingly, in this quadratic example, the best disclosure policy for the Principal is full transparency no matter how big b, the reason being that anyway the agent will apply the bias b to his choice of a and given the concavity of $a \to -(a-\theta)^2$ it is never in the interest of the Principal that the agent be incompletely informed of θ.[7]

3.2 Basic considerations

The Crawford-Sobel quadratic example casts doubt as to whether the misalignment of preferences of the principal and the agent is the only driving force behind the desirability of non-transparency. In this subsection, I suggest refining the required notion of misalignment by considering two forces that relate to how the concavity of the utility function of the agent and the sensitivity of the Principal's objective to the agent's action vary with the state. To formulate this, I consider the case of the bundling of two states θ_1 and θ_2 into one information set $I = \{\theta_1, \theta_2\}$ for the agent, and I illustrate two extreme cases in which the Principal may prefer this to letting the agent know whether $\theta = \theta_1$ or θ_2. The first of these cases is related to the relative concavity of the agent's utility function with respect to his own action at these two states while the second of these cases is related to the relative sensitivity of the principal's utility function with respect to the agent's action. Both will play a key role in understanding the general non-transparency insight to be derived in Section 4.[8]

To present this most simply, assume that the action space A is one-dimensional, and let

$$\begin{aligned} a(\theta) &= \arg\max_a u(a;\theta) \\ a^{CI} &= \arg\max_a [p(\theta_1)u(a;\theta_1) + p(\theta_2)u(a;\theta_2)] \end{aligned}$$

[7] When several θ belong to the same information set I, the agent picks $a^{CI} = E(\theta \mid \theta \in I) + b$ and conditional on $\theta \in I$ the Principal gets $-E[(a^{CI}-\theta)^2 \mid \theta \in I]$ which is less (by Jensen's inequality) than $-(a^{CI} - E(\theta \mid \theta \in I))^2 = -b^2$. Given that $-b^2$ is what the Principal gets under full transparency, we get the desired conclusion that transparency is optimal whatever b (and whatever the distribution of θ).

[8] A third effect in favor of non-transparency one might think of is related to the concavity of the principal's payoff function π with respect to the agent's action a, but this turns out not to play any role in the derivation of the non-transparency result, as will be explained next.

Assume further that $a(\theta_2) > a(\theta_1)$. The concavity of $a \rightarrow u(a;\theta)$ for all θ ensures that

$$a(\theta_1) < a^{CI} < a(\theta_2)$$

Relative concavity of u

Clearly, if $u(\cdot;\theta)$ is much more concave at θ_2 than at θ_1 -in the sense that $\left|\frac{\partial^2 u}{\partial a^2}(\cdot;\theta_2)\right|$ is sufficiently (and uniformly) bigger than$\left|\frac{\partial^2 u}{\partial a^2}(\cdot;\theta_1)\right|$- then a^I will be close to $a(\theta_2)$ and to the extent that $\pi(\cdot;\theta)$ is increasing in a, non-transparency will be preferred to transparency.

Relative sensitivity of π

For a given $u(\cdot;\theta)$, if the principal cares much more about the agent's effort a at θ_1 than at θ_2 -in the sense that $\left|\frac{\partial \pi}{\partial a}(\cdot;\theta_1)\right|$ is sufficiently (and uniformly) bigger than$\left|\frac{\partial \pi}{\partial a}(\cdot;\theta_2)\right|$- then the increase of a (from $a(\theta_1)$ to a^I) at θ_1 will more than compensate (in terms of Principal's payoff) the decrease of a (from $a(\theta_2)$ to a^I) at θ_2 so that non-transparency is preferred to transparency.

3.3 An illustrative example

In order to fix the kind of applications I have in mind, consider the following moral hazard scenario. An agent must choose an effort level $a \in [0,1]$. The cost of exerting effort a is $c(a) = \frac{1}{2}a^2$. Depending on the effort a, the probability of success is γa where $\gamma < 1$. Success is not directly observed by the Principal (because there are many activities in the organization and it is not obvious to pin down whether an isolated activity is successful or not). What the Principal observes is a signal σ which can take two values $\sigma = 1$ or 0. The observation of σ provides a noisy signal as to whether the activity is successful. Specifically, assume that $\Pr(\sigma = 1 \mid success) = \beta > \frac{1}{2}$. The agent receives a bonus of 1 in case $\sigma = 1$ is observed. The Principal receives an extra benefit R in case of success. Both the Principal and the agent are assumed to be risk neutral.

The state that describes the interaction can be parameterized by (β, γ, R). To simplify, assume that R does not vary so that the state can be reduced to $\theta = (\beta, \gamma)$ where

β parameterizes the monitoring technology of the Principal and γ parameterizes how difficult the task is.

In a world in which the agent would know the state θ, he would pick action $a(\beta, \gamma)$ so as to maximize his expected utility $[\gamma a\beta + (1 - \gamma a)(1 - \beta)] - \frac{1}{2}a^2$, where the first (resp second) term in the bracket represents the probability that the Principal observes $\sigma = 1$ and the task is successful (resp. non-successful) and thus the overall term in bracket represents the expected bonus received by the agent when exerting effort a. Optimality with respect to a requires that:

$$a(\theta) = \gamma(2\beta - 1)$$

When the agent does not know the state θ, he would pick

$$a^{CI} = E[\gamma(2\beta - 1)]$$

where E denotes the expectation operator over realizations of $\theta = (\beta, \gamma)$.

The Principal's payoff is the reward R in case of success diminished by the bonus in case $\sigma = 1$ is observed. The principal prefers the non-transparent scenario whenever the expectation (over θ) of

$$\gamma a^{CI}[R - (2\beta - 1)] - (1 - \beta)$$

is greater than the expectation of

$$\gamma a(\theta)[R - (2\beta - 1)] - (1 - \beta)$$

That is, whenever

$$RE(\gamma)E(\gamma(2\beta - 1)) - [E(\gamma(2\beta - 1))]^2 > RE(\gamma^2(2\beta - 1)) - E(\gamma^2(2\beta - 1)^2)$$

To simplify a bit, consider the special case in which R would be very large so that the superiority of non-transparency boils down to whether the following condition holds

$$E(\gamma)E(\gamma(2\beta - 1)) > E(\gamma^2(2\beta - 1)). \tag{1}$$

This condition says that it is best for the Principal not to let the agent know the state whenever γ and $\gamma(2\beta-1)$ are negatively correlated. Just to illustrate how much can be gained with non-transparency, assume that there are two states equally likely. The monitoring technology takes two possible forms: Either the signal σ received by the Principal is totally uninformative about success so that $\beta = \frac{1}{2}$, or the signal σ is perfectly informative so that $\beta = 1$. Moreover, when the signal is uninformative about success, the sensitivity of success to effort is highest with $\gamma = 1$, and when the signal σ is perfectly informative about success, the sensitivity of success to effort is lower and $\gamma = \underline{\gamma} < 1$. In such a case, non-transparency gives an expected payoff to the Principal proportional to $(1+\underline{\gamma})\underline{\gamma}$, whereas transparency gives an expected payoff proportional (with the same coefficient) to $(\underline{\gamma})^2$. The gain of non-transparency in relative terms is then $\frac{1-\underline{\gamma}}{2\underline{\gamma}}$ and it can be made arbitrarily large as $\underline{\gamma}$ is made smaller and smaller. Here non-transparency is good because it allows to incentivize the agent to exert more effort in the state in which effort induces more success.

Observe that in the two state scenario just considered if γ and $\gamma(2\beta-1)$ were positively (instead of negatively) correlated, then transparency over these two states would be better than non-transparency. Yet, as I show later, in cases like the one considered here in which the same action $a(\theta)$ would be chosen at several states were the agent to be fully informed of the state, one can always find a set of states so that if the agent is not told which state in the set occurred, the Principal is strictly better off than if the agent is fully informed of the state.

4 Non-transparency

In this Section, I establish quite generally that if the same action a were to be chosen in at least two distinct states θ_1, θ_2 in the benchmark scenario in which the agent would be fully informed of the state, then for generic objective functions π, the designer can achieve a strictly higher expected payoff by not disclosing fully θ. Thus, some form of non-transparency is desirable. As a corollary, I obtain that full transparency is not the optimal disclosure policy whenever the dimension of the state space exceeds the dimension of the action space because in such cases there must exist distinct states that induce the same choice of action in the complete information benchmark.

4.1 The basic non-transparency result

To establish my first non-transparency result, let me denote by $a(\theta) = \arg\max_a u(a;\theta)$. I will assume that for all θ, there is always a direction δ of θ such that

$$\frac{\partial}{\partial \delta} \nabla_a u(a(\theta);\theta) \neq 0$$

where $\nabla_a u$ refers to the gradient of u with respect to a. When u satisfies this assumption, I say that u is non-satiated. Such an assumption typically implies that maxima $a(\theta)$ are never locally constant (it is in fact equivalent to this). It should typically be thought of as a fairly weak assumption as such a direction of θ may for example correspond to a reduction in the marginal cost of producing effort in some of the tasks to be fulfilled by the agent. This assumption was trivially met in the example considered in Section 3.

I will also need to define more formally what I mean by generic objective functions π. I provide below a set theoretic definition of genericity (so as to deal with the space of objective functions in great generality) but note that a measure theoretic definition would work equally well if I were to consider that objective functions can only take polynomial forms, say. Formally, let $X = \mathbb{R}^n \times \mathbb{R}^m$ denote the domain of the objective functions π. Consider functions $\pi \in C^2(X)$. The set $\overline{\Pi}$ of $\pi \in C^2(X)$ is endowed with a Whitney C^2 topology by letting a sequence $\pi_k \in \overline{\Pi}$ converge to π if and only if $\pi_k - \pi$ as well as the Jacobian of $\pi_k - \pi$ and the matrix of second derivative of $\pi_k - \pi$ converge uniformly to zero in the space of continuous functions over the relevant range of a, θ (assumed to be bounded, see above). Genericity is defined as:

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 first main result is:

Proposition 1 *Assume that there exist two distinct states θ_1, θ_2 in the interior of the θ-space such that the same action would be chosen in the complete information benchmark, i.e. $a(\theta_1) = a(\theta_2)$. Assume that the utility function u is non- satiated. There exists a* *generic set Π such that for all objective functions $\pi \in \Pi$, the Principal can do strictly better than disclosing fully the state θ to the agent.*

The general case covered by Proposition 1 is shown in Appendix. I now provide a detailed intuition for the case in which the action space is one-dimensional $a \in \mathbb{R}$. The complete information solution $a(\theta)$ satisfies: $\frac{\partial}{\partial a}u(a(\theta);\theta) = 0$.[9] Let θ_1 and $\theta_2 \neq \theta_1$ be such that $a(\theta_2) = a(\theta_1)$, as allowed by the Proposition, and consider a direction δ such that $\frac{\partial^2 u}{\partial a \partial \delta}(a(\theta_1);\theta_2) \neq 0$, as allowed by the non-satiation of u.

Consider the states $\theta = \theta_1$ and $\theta_2 + \varepsilon\delta$ for ε either positive or negative but small (remember θ_2 lies in the interior of the θ-space). The central part of the argument consists in comparing the aggregate expected organizational payoff π when the agent knows whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ and the expected organizational payoff when the agent ignores whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$. From there, it will be relatively straightforward that the Principal can strictly gain in expected terms as compared with the full information benchmark.

Clearly for $\varepsilon = 0$, the two informational scenarios generate the same aggregate expected value of π. 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 benchmark.

Let $a_1 = a(\theta_1)$ and $a_2(\varepsilon) = a(\theta_2 + \varepsilon\delta)$. They satisfy

$$\begin{aligned} \frac{\partial}{\partial a}u(a_1;\theta_1) &= 0 \\ \frac{\partial}{\partial a}u(a_2(\varepsilon);\theta_2 + \varepsilon\delta) &= 0 \end{aligned} \tag{2}$$

Let $a^{CI}(\varepsilon)$ denote the action when the agent does not know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$.

[9] Observe that what I am assuming here is only that (at least over a range of α) the solution $a(\theta)$ varies smoothly with θ and is locally pinned down by the first-order conditions (i.e., no other a in the neighborhood of $a(\theta)$ satisfies the first-order condition).

It satisfies:

$$p(\theta_1)\frac{\partial}{\partial a}u(a^{CI}(\varepsilon);\theta_1)+p(\theta_2+\varepsilon\delta)\frac{\partial}{\partial a}u(a^{CI}(\varepsilon);\theta_2+\varepsilon\delta)=0. \tag{3}$$

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

$$p(\theta_1)[\pi(a_1;\theta_1)-\pi(a^{CI}(\varepsilon);\theta_1)]+p(\theta_2+\varepsilon\delta)[\pi(a_2(\varepsilon);\theta_2+\varepsilon\delta)-\pi(a^{CI}(\varepsilon);\theta_2+\varepsilon\delta)].$$

Clearly, if $\Delta(\varepsilon)<0$, it is strictly better that the agent does not know whether $\theta=\theta_1$ or $\theta_2+\varepsilon\delta$.

I now expand $\Delta(\varepsilon)$ at the first order in ε. Since $a^{CI}(0)=a_2(0)=a_1$, $\Delta(\varepsilon)$ writes at the first order:

$$p(\theta_1)\frac{\partial\pi}{\partial a}(a_1;\theta_1)[a_1-a^{CI}(\varepsilon)]+p(\theta_2)\frac{\partial\pi}{\partial a}(a_1;\theta_2)[a_2(\varepsilon)-a^{CI}(\varepsilon)]+o(\varepsilon)$$

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

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

$$\begin{aligned} a_2(\varepsilon)-a_1 &= \frac{-\frac{\partial^2 u}{\partial a\partial\delta}(\theta_2)}{\frac{\partial^2 u}{\partial a^2}(\theta_2)}\varepsilon+o(\varepsilon)\\ a^{CI}(\varepsilon)-a_0 &= \frac{-p(\theta_2)\frac{\partial^2 u}{\partial a\partial\delta}(\theta_2)}{p(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_1)+p(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_2)}\varepsilon+o(\varepsilon)\end{aligned}$$

where $\partial h/\partial\delta$ denotes the derivative of (an arbitrary function denoted) h in the direction δ and all functions are taken at $a=a_1$.

After multiplying $\Delta(\varepsilon)$ by $\frac{\partial^2 u}{\partial a^2}(\theta_2)[p(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_1)+p(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_2)]$ and dividing by $p(\theta_1)p(\theta_2)$ (which are both strictly positive) we get that $\Delta(\varepsilon)$ has the same sign as

$$\left[\frac{\partial\pi}{\partial a}(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_2)-\frac{\partial\pi}{\partial a}(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_1)\right]\frac{\partial^2 u}{\partial a\partial\theta^\delta}(\theta_2)\varepsilon+o(\varepsilon)$$

Three cases may a priori occur.

1) $\left[\frac{\partial\pi}{\partial a}(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_2)-\frac{\partial\pi}{\partial a}(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_1)\right]\frac{\partial^2 u}{\partial a\partial\delta}(\theta_2)<0$. Then taking $\varepsilon>0$ and sufficiently small, I can infer from the above that not letting the agent know whether $\theta=\theta_1$ or

$\theta_2 + \varepsilon\delta$ strictly dominates the complete information benchmark.

2) Likewise, if $\left[\frac{\partial \pi}{\partial a}(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_2) - \frac{\partial \pi}{\partial a}(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_1)\right]\frac{\partial^2 u}{\partial a \partial \delta}(\theta_2) > 0$, then taking $\varepsilon < 0$ and sufficiently small, not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ strictly dominates the complete information benchmark (remember than since θ_2 is in the interior of the θ-space, one can move in any direction from θ_2).

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

$$\left[\frac{\partial \pi}{\partial a}(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_2) - \frac{\partial \pi}{\partial a}(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_1)\right]\frac{\partial^2 u}{\partial a \partial \delta}(\theta_2) = 0$$

or

$$\frac{\partial \pi}{\partial a}(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_2) - \frac{\partial \pi}{\partial a}(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_1) = 0 \tag{4}$$

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

To see this formally, consider the family of π_λ functions

$$\pi_\lambda(a;\theta) = \pi(a;\theta) + \lambda a \parallel \theta - \theta_1 \parallel^2$$

where $\lambda \in \mathbb{R}$ and $\parallel \theta - \theta_1 \parallel$ denotes the euclidean distance between θ and θ_1. Obviously, if π satisfies (4), then for $\lambda \neq 0$, π_λ does not satisfy (4) -observe that changing π does not affect the expressions of $a_2(\varepsilon)$, $a^{CI}(\varepsilon)$- from which one can conclude that the set of π for which (4) does not hold is dense. Moreover, this set is also open given the continuity of the mapping $\pi \rightarrow \frac{\partial \pi}{\partial a}(\theta_1)\frac{\partial^2 u}{\partial a^2}(\theta_2) - \frac{\partial \pi}{\partial a}(\theta_2)\frac{\partial^2 u}{\partial a^2}(\theta_1)$ according to the Whitney C^2 topology.[10]

The rest of the argument, in particular showing how one can induce a strict gain in expectation (integrating over all possible realizations of θ) relies on a simple continuity argument. If not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ is strictly preferable to full transparency (as just considered) then there must exist (small enough) neighborhoods B_1 , B_2 of θ_1 and $\theta_2 + \varepsilon\delta$ respectively, such that not letting the agent know $\theta \in B_1 \cup B_2$ induces an expected (strictly) positive gain in terms of the Principal's objective as compared with the full information benchmark. Observe that this gain can be achieved even in the scenario in which the Principal is constrained either to disclose the state θ or

[10] Clearly, if one were to consider polynomial functions π, then (4) would not hold for a measure 1 set of parameter values.

to remain silent (as considered later in an application).

4.2 Discussion

In this subsection, I first interpret the two-state construction shown above in light of the basic considerations made in Section 3. I then show the implication of Proposition 1 for the case in which the dimension of the state space is higher than the dimension of the action space.

4.2.1 Why is full transparency dominated?

Suppose in the two-state scenario considered above that the action of the agent is strictly larger at $\theta_2 + \varepsilon\delta$ than at θ_1, that is, $a(\theta_2 + \varepsilon\delta) > a(\theta_1)$.

If the Principal is more sensitive to effort at θ_1 than at θ_2 in the sense that $\frac{\partial \pi}{\partial a}(\theta_1) > \frac{\partial \pi}{\partial a}(\theta_2)$ while the agent's utility function has the same concavity at the two states ($\frac{\partial^2 u}{\partial a^2}(\theta_1) = \frac{\partial^2 u}{\partial a^2}(\theta_2)$), then not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ is good for the Principal because it allows to transfer some effort made in state $\theta_2 + \varepsilon\delta$ to state θ_1 and the Principal is more sensitive to extra effort in state θ_1.

Symmetrically, if the Principal is equally sensitive to the agent's effort in states θ_1 and θ_2 ($\frac{\partial \pi}{\partial a}(\theta_1) = \frac{\partial \pi}{\partial a}(\theta_2) > 0$) and the agent's utility function is more concave at θ_1 than at θ_2 in the sense that $\mid \frac{\partial^2 u}{\partial a^2}(\theta_1) \mid > \mid \frac{\partial^2 u}{\partial a^2}(\theta_2) \mid$ then not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ is good for the Principal because the effort made in the coarse information case is closer to that made in state $\theta_2 + \varepsilon\delta$ than to that made in state θ_1, which is assumed to be larger.

The above two cases for non-transparency are, of course, closely related to the basic observations made in Section 3. Incorporating both effects of the relative sensitivity of π to a on the one hand and the relative concavity of u on the other reveals that the coarse information structure dominates full transparency whenever

$$-\frac{\partial \pi / \partial a}{\partial^2 u / \partial a^2}(\theta_1) > -\frac{\partial \pi / \partial a}{\partial^2 u / \partial a^2}(\theta_2) \qquad (5)$$

which provides an exact expression of how the two effects should be aggregated.

Observe that the key observation behind Proposition 1 is that either (5) is satisfied and not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ is good for the Principal or else we should generically have that $-\frac{\partial\pi/\partial a}{\partial^2 u/\partial a^2}(\theta_1) < -\frac{\partial\pi/\partial a}{\partial^2 u/\partial a^2}(\theta_2)$ and then not letting the agent know whether $\theta = \theta_1$ or $\theta_2 - \varepsilon\delta$ is good for the Principal.

Two remarks may be worth making here. First, it should be noted that the argument developed above makes no use of the concavity of the objective function π with respect to a. One might have conjectured that the concavity of π would play a role in the argument given that a coarser information of the agent reduces the variability of the action with the state and this reduced variability would seem to be favorable to the Principal whenever π is concave in a. Yet, as Proposition 1 implies, the desirability of some form of non-transparency holds no matter whether π is concave or not, in particular even in cases in which π would be extremely convex.

The reason for this seemingly surprising result is that by considering the bundling of the states $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ which are such that $a(\theta_1) = a(\theta_2)$, the concavity of π has an effect of order ε^2 on the comparison as to whether the Principal prefers that the agent be uninformed about the state (by contrast, the previous highlighted effects are of order ε). As a result, the concavity of π plays no role as to whether some form of non-transparency may be desirable under the assumptions made in Proposition 1. Another way to phrase this is that the choice of information structure made above is such that the effect of the concavity of π is frozen and only the relative concavity of u and the relative sensitivity of π with respect to a play a role. Of course, the concavity of π with respect to a will play a role as to what exact form of information disclosure is best for the Principal. But no matter what π looks like, some form of non-transparency will be desirable, as long as there exist at least two distinct states that would induce the same choice of action in the complete information benchmark.[11]

It should also be stressed that if in the above construction, we had considered $\theta_2 = \theta_1$, then the first-order effects as expressed above would cancel out (this can be seen by

[11] Maybe in relation to the latter point, it can be said that if we had considered two distinct states θ_1, θ_2 with $a(\theta_1) \neq a(\theta_2)$, then the concavity of π would have played a non-negligible role in the assessment as to whether the Principal would be better off when the agent does not know the state $\theta \in \{\theta_1, \theta_2\}$. More precisely, it is readily verified that, for sufficiently convex π, the Principal would have been strictly better off in the full information benchmark as compared with the situation in which the agent does know whether $\theta = \theta_1$ or θ_2.

plugging $\theta_2 = \theta_1$ in the above expressions), and thus whether the Principal is better off when the agent knows whether $\theta = \theta_1$ or $\theta_1 + \varepsilon\delta$ would involve a comparison of order ε^2 and it would thus not be possible to conclude that some form of non-transparency is desirable (considering the bundling of θ_1 or $\theta_1 - \varepsilon\delta$ as compared with the fully transparent benchmark would now have an effect of the same sign as the bundling of θ_1 or $\theta_1 + \varepsilon\delta$, and as a matter of fact, it may well be then that full transparency dominates as illustrated in the quadratic example of Crawford-Sobel).

Comment. As illustrated above, in the case in which the dimension of the action space is 1, the relative density $p(\theta)$ around $\theta = \theta_1$ and θ_2 plays no role as to whether the Principal is better off when the agent does know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$. However, when the action space has a dimension larger than 1, this relative density may play a role. Inspecting the proof of Proposition 1 in the Appendix reveals this.

4.2.2 When the state space has more dimensions than the action space

A simple corollary of Proposition 1 is obtained when the state $\theta \in \mathbb{R}^m$ varies over more dimensions than the action $a \in \mathbb{R}^n$, i.e $m > n$. Indeed in such a case, there are typically many different states θ that induce the same action in the full information benchmark. Specifically, given the smoothness of u, for almost any θ_1,

$$\overline{\Theta}(\theta_1) = \{\theta \text{ such that } a(\theta) = a(\theta_1)\}$$

defines a manifold of dimension $m - n$ (see, for example, Milnor (1965), chapter 2). Thus one can find $\theta_2 \neq \theta_1$ such that $a(\theta_2) = a(\theta_1)$. As a consequence of Proposition 1, I can state:

Theorem 1 *Suppose the dimensionality of the state θ is strictly bigger than the dimensionality of the action a of the agent, that is, $m > n$, and suppose the utility of the agent is non-satiated. Then there exists a generic set Π such that for all objective functions $\pi \in \Pi$, some non-full disclosure of θ to the agent is strictly beneficial to the Principal as compared with the full information benchmark.*

I think that the scope of application of Theorem 1 is significant given that in many applications it would seem reasonable to believe that the state varies over more dimensions than the action. In the leading example of Section 3, the state included a description of how hard the task was and what kind of signals was being observed by the Principal whereas the action took the form of an effort level, thereby satisfying the condition of Theorem 1. Moving into the territory of multi-tasking in which the dimension of the action space would be greater, one should have in mind that the type of signals that the Principal could observe would then be of greater richness and also that the state could include descriptions of how complementary the tasks of the agent are to the working of the organization (about which the Principal could get noisy signals), thereby suggesting that $m > n$ is still a good assumption in applications of this sort.

In some applications, it may be hard not to let the agent know about some aspects of the state. For example, if the state includes a description of the private costs incurred by the agent when exerting a certain type of effort, it would seem natural to assume that the agent knows about this no matter what disclosure policy is chosen by the Principal. With this constraint in mind, one should reinterpret Theorem 1 as follows: Some form of non-transparency is desirable, as long as the number of dimensions of the state that can be hidden to the agent is strictly larger than the number of dimensions of the action to be chosen by the agent. I believe such a dimensional gap between what can be hidden from the state and the action of the agent is quite common in applications, and the leading example of the paper provides a natural illustration of this.

Comment. To illustrate why the non-transparency conclusion of Theorem 1 does not hold whenever the action space and the state space have the same dimension, one can specialize the above moral hazard example and assume that the monitoring technology as measured by β does not vary, so that the state is fully parameterized by $\theta = \gamma$. In this case, by Jensen's inequality (and the convexity of $x \to x^2$) it is readily verified that (1) does not hold so that full transparency is the best disclosure policy. The desirability of transparency in this case should sound economically intuitive. If the only heterogeneity lies in the sensitivity γ of success to effort, the Principal would like the highest effort to be made when success is easier to achieve, which is most efficiently obtained when the agent knows more about the state.[12]

5 Extensions

In this Section, I establish the robustness of the non-transparency insight derived in Section 4 in two directions. First, I consider the case of several agents. Second, I allow for the possibility of using optimally designed monetary instruments. In both extensions, I observe that the non-transparency insight holds, as long as the dimension of what can be hidden to the agent(s) exceeds the dimension of the action(s). Thanks to these extensions, I believe the scope of application of the main non-transparency insight of this paper is quite large.

5.1 Multi-agent interactions

Compared to the model described in Section 2, I allow the presence of several agents. Each agent i must now simultaneously choose an action $a_i \in A_i$ where A_i is an open subset of $\mathbb{R}^{n_i}$. The state is still denoted by θ and it belongs to $\mathbb{R}^m$. When the profile of actions is $a = (a_i)_i$ and the state is θ, the utility derived by agent i is denoted by $u_i(a;\theta)$, and the payoff to the Principal is denoted by $\pi(a;\theta)$. All utility functions are assumed to be smooth functions of the relevant variables, and for each agent i, u_i is assumed to be concave in a_i.

In the fully transparent benchmark, for each state θ, the play of agents is described by a Nash equilibrium of the corresponding normal form game. Let me denote by $a^{NE}(\theta) = (a_i^{NE}(\theta))_i$ the action profile that would be played in state θ. I will assume that in some open neighborhood of states, $a^{NE}(\theta)$ varies smoothly with θ, and in the rest of this Section

[12] It should be mentioned that in the context of this example, if the sole heterogeneity lies in the monitoring technology, then the Principal is better off hiding it to the agent (this follows again from Jensen's inequality and the convexity of $x \to x^2$). Yet, if one were to enrich the example and allow the Principal to enjoy a benefit R that is sufficiently increasing in the agent's effort a, then full transparency would dominate in the case the sole heterogeneity lies in the monitoring technology β.

I am assuming the state space is included in such a neighborhood.[13]

As in Section 3.1, I will assume in the following Proposition that there are two distinct states θ_1 and θ_2 such that $a_i^{NE}(\theta_1) = a_i^{NE}(\theta_2)$ for some agent i. Consider the Bayesian game in which agent i does not know whether $\theta = \theta_1$ or θ_2, while agents other than i know the state. One Bayes Nash equilibrium of this Bayesian game would result in the same distribution of action profiles as in the complete information benchmark. Also, thanks to our smoothness assumptions, if agent i is uninformed as to whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ for some direction δ and some scalar ε, there will exist a Nash Bayes equilibrium of the corresponding Bayesian equilibrium that varies smoothly with ε and such that it converges to the complete information Nash equilibrium a^{NE} as ε goes to 0. When considering Bayesian games, I make the assumption that the play follows such a Bayes Nash equilibrium.

Proposition 2 *Assume that there exist two distinct states θ_1, θ_2 in the interior of the θ-space such that the same action would be chosen by agent i in the complete information benchmark, i.e. $a_i^{NE}(\theta_1) = a_i^{NE}(\theta_2)$, and that the utility function u_i of agent i is non-satiated.*[14] *There exists a generic set Π such that for all objective functions $\pi \in \Pi$, the Principal can do strictly better than disclosing fully the state θ to the agents.*

The argument to prove Proposition 2 is very similar to that used to prove Proposition 1. Consider the incomplete information setting in which agents other than i know the state, and agent i does not know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ where δ is a direction such that $\frac{\partial}{\partial\theta^\delta}\nabla_{a_i} u_i(a^{NE}(\theta_2); \theta_2) \neq 0$. The key step consists in showing that the difference of Principal's expected payoff in this incomplete information setting and in the complete information benchmark is of the same order as ε for generic π functions, and thus either $\varepsilon > 0$ or $\varepsilon < 0$ but small would ensure the strict superiority of the incomplete information scenario. The argument is detailed in the Appendix.

[13] For every θ, an equilibrium in pure strategies exists thanks to our concavity assumption. From the Nash equilibrium correspondence, I am considering here a selection that varies smoothly with θ (which can be done thanks to our smoothness assumptions at least locally using again the techniques of topology from the differentiable viewpoint, Milnor 1965).

[14] Non-satiation is defined here with respect to a_i, i.e., $\forall\theta$, $\exists\delta$ such that $\frac{\partial}{\partial\delta}\nabla_{a_i} u_i(a^{NE}(\theta); \theta) \neq 0$.

As in the one agent case, if the dimension m of the state θ is strictly bigger than the dimension n_i of the action a_i of at least one agent i, then the conditions for Proposition 2 will be satisfied for this agent, and thus some form of non-transparency will be desirable.

In the arguments just developed, I have allowed the Principal to use a different disclosure policy for the various agents. If the disclosure policy must be the same for all agents, then the same conclusion as that in Proposition 2 can be reached, as long as there exist two distinct states θ_1, θ_2 such that the same action profile would be chosen at the two states in the complete information benchmark, i.e. $a^{NE}(\theta_1) = a^{NE}(\theta_2)$. As a corollary, I can infer that some form of non-transparency is desirable in the public information disclosure case, as long as the dimension m of θ is strictly larger than the sum $\Sigma_i n_i$ of dimensions of a_i over all agents i.

5.2 Monetary instruments

Monetary instruments can be used in several ways. First, one can think of the Principal who possesses private information on θ as being able to sell her information to the agent. The difference with the previous analysis is that now the Principal is not reduced to disclose freely some aspects of her information. She can disclose this in exchange for side-payments. Second, in the tradition of contract theory, monetary instruments can be used to better align the incentives of the agent with the interest of the Principal making the side-payments contingent on what the Principal observes ex post. In both cases, it is of interest to review what happens to the main non-transparency result derived above.

5.2.1 The sale of information

A simple way to model this is to assume that some bargaining between the Principal and the agent takes place at the ex ante stage before the realization of θ is known where bargaining bears on which disclosure policy the Principal will commit to in exchange for side-payments.

It is not difficult to see that the non-transparency result as described in Proposition 1 or Theorem 1 continues to hold in such a setting no matter what the exact details of the bargaining protocol are.

To see this most simply, assume that the utilities of the Principal and the agent are quasi-linear in money (where the dependence of u and π in the agent's action a are still assumed to be of the most general form). Then, bargaining necessarily results in the maximization of $u + \pi$, and it is not difficult to infer from Proposition 1 that, as long as the utility function u of the agent is non-satiated and there are two distinct states θ_1, θ_2 such that the same action would be chosen in the complete information benchmark, i.e. $a(\theta_1) = a(\theta_2)$, then full transparency is not the chosen disclosure policy for generic objective functions π (because genericity of $\pi + u$ as confined to the set of preferences that are quasi-linear in money immediately translates into genericity of π viewed as a function of (a, θ)). The same conclusion would also hold true in the case of non-quasi-linear utilities even though the argument would have now to make more precise the bargaining protocol (as affecting say the bargaining power of the two parties) as it may now have an impact on which disclosure policy would be adopted.

Despite its simplicity, the argument above should be contrasted with the result obtained by Eso and Szentes (2007) showing that it is optimal for the seller of a good to disclose as much as she can to potential buyers in a context of private value auctions with quasi-linear preferences. Assuming away the private information of the buyers, it is readily verified that in the context studied by Eso and Szentes, the welfare-maximizing outcome can be achieved whenever the seller discloses all what she knows, thereby explaining Eso-Szentes' insight in terms of the optimal disclosure policy. Yet, in general moral hazard problems, even assuming utilities are quasi-linear in money, there is no reason why full disclosure would result in the maximization of total welfare, and thus the full transparency insight of Eso and Szentes does not extend to such contexts.

5.2.2 Side-payments as incentive instruments

A full description of the state θ should now include a specification of what the Principal observes ex post and how this is affected by the action a of the agent. To simplify, assume that the Principal can only observe a finite number s of signals $\sigma = 1, ...s$. Then the monetary instrument available to the Principal can be described as a vector $w = (w_\sigma)_\sigma \in \mathbb{R}^s_+$ specifying the side-payment assumed to be non-negative or bonus $w_\sigma \in \mathbb{R}_+$ the agent would receive from the Principal in case the Principal observes σ ex post. In state θ, action a would result in a probability distribution over σ that I denote by $q_\theta(\cdot \mid a)$, and I assume that for all θ, σ and a, $q_\theta(\sigma \mid a) > 0$.[15] Agent's expected payoff is $u(a; w, \theta)$ when w prevails, the state is θ, and the agent chooses a. The corresponding expected payoff to the Principal is denoted by $\pi(a; w, \theta)$, and it is assumed to be decreasing in w. All functions are assumed to vary smoothly with a, w, θ. Besides u is assumed to be a concave function of a for all w and θ. Moreover, I assume that the function $a \to u(a; w, \theta)$ is always maximized in a bounded subset of A,[16] and I denote the maximand by $a(w, \theta)$. Finally, I denote by $w(\theta)$ the bonus scheme that maximizes $\pi(a(w, \theta); w, \theta)$.

I rephrase the transparency question for the case with monetary transfers as follows:

Question. Can it be beneficial for the designer that the agent be partially rather than fully informed of (θ, w)?

Compared to the case without monetary transfers, transparency is defined as a situation in which the agent knows the state θ as well as the exact incentive scheme as defined by w that prevails in θ. Observe that if the Principal were to choose a stochastic incentive scheme (a distribution over w) and have the agent know the state θ (but not the realization of w), this would be considered a form of non-transparency in the context of the above question (even if I conjecture that a non-transparency result solely in terms of partial disclosure of the state θ is likely to hold even in cases in which stochastic contracts are optimal).

The notion of genericity is the same as that introduced in Section 4 except that the domain of π is now $X = \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^m$ given that π depends now also on the monetary instruments $w \in \mathbb{R}^s$. We have:

Proposition 3 *Assume that there exist two distinct states θ_1, θ_2 in the interior of the θ-space such that the same action would be chosen in the complete information benchmark,*

[15] This full support assumption typically implies that the first-best cannot be achieved in the full information benchmark

[16] Such assumptions will typically guarantee that one can work with first-order conditions to deal with the maximization problems of agents. To the extent that interior pure strategy equilibria exist and that one focuses on these, the insights developed below would apply.

i.e. $a(w(\theta_1), \theta_1) = a(w(\theta_2), \theta_2)$*, and that the utility function* u *is non- satiated. There exists a generic set* Π *such that for all objective functions* $\pi \in \Pi$*, the Principal can do strictly better than disclosing fully the state* θ *to the agent.*

The idea of the proof is very similar to that of Proposition 1. For each θ, define $\widetilde{u}(a; \theta) = u(a, w(\theta); \theta)$, and $\widetilde{\pi}(a; \theta) = \pi(a, w(\theta); \theta)$. The genericity of π translates into the genericity of $\widetilde{\pi}$, and one can apply Proposition 1 to derive Proposition 2. The economic intuition for this result is that even if the monetary instrument w is optimally adjusted to the state θ, $u(a, w; \theta)$ is not perfectly aligned with $\pi(a, w; w)$ and thus the Principal can gain by using clever non-transparent disclosure policies.[17]

Of course, similarly to Theorem 1, I also observe that as soon as the dimension of the state space m is larger than the dimension of the action space n, i.e. $m > n$, one must have two distinct states θ_1 and θ_2 such that $a(w(\theta_1), \theta_1) = a(w(\theta_2), \theta_2)$ and thus some form of non-transparency is desirable.

Discussion

1) The above result about the desirability of some form of non-transparency in the presence of monetary instruments seems to be at odds with the conclusion obtained by Rayo and Segal (2010) who show in their setup in which the state space has dimension 2 and the action space has dimension 1 that when monetary instruments are allowed, the best disclosure policy is full transparency. Their model can be described within my setup as follows. There is one agent choosing an action $a \in [0, 1]$ (called an acceptance probability in Rayo and Segal). The monetary instrument (when available) takes the form of a transfer w from the principal to the agent in case of acceptance. The state

[17] Based on this intuition, I would expect the result of Proposition 3 to continue to hold even if the observation made by the Principal were to take values on the continuum, but the formal argument would require dealing with First Order Conditions in the space of bonus functions which would be technically more challenging (than just dealing with local optima over finitely many bonuses for each possible observation).

$\theta = (\tau, \phi) \in \mathbb{R}^2$ is two-dimensional and affects payoffs as follows:

$$\begin{aligned} \pi(a;\theta,w) &= \tau a - wa \\ u(a;\theta,w) &= -\frac{1}{2}a^2 + \phi a + wa \end{aligned}$$

The main reason why full transparency is best when w can be made contingent on the state θ is that with the chosen parametric forms the Principal can always replicate any outcome obtained with partial disclosure of θ using well chosen monetary instruments and full disclosure, and still obtain the same expected payoff. It follows that full transparency cannot be worse than any alternative disclosure policy, and thus it is a best policy.[18] Yet, such a property can only hold for special (non-generic) specifications of π, and in general some form of non-transparency would be desirable even in the case in which side-payments are allowed.

2) The general issue as to whether transparency is desirable can be formulated in terms of the following trade-off. If an agent receives incomplete information about the state θ, it makes it easier to satisfy the incentive constraints of the agent through an appropriate use of the monetary instruments. This is because incomplete information allows the designer to aggregate the various incentive constraints (for each state) into one incentive constraint that is easier to satisfy. But, incomplete information forces the agent not to be able to adjust his action to the state θ, which is sometimes harmful to the designer. In light of this, what Proposition 3 establishes is that one can always find information structures in which the first effect dominates the second, as soon as there are at least two distinct

[18] To see this, consider the case in which the agent is uninformed of θ (within some set) and let $w(\theta)$ be the monetary instrument adopted by the principal when θ prevails. The agent chooses:

$$a^{CI} = E_\theta(\phi + w(\theta))$$

and the designer gets an expected payoff equal to $E(\tau)a^{CI} - E(w(\theta))a^{CI}$.

Let us now consider the case in which the agent is fully informed of θ and the monetary scheme $w^*(\theta)$ is defined so that:

$$a^{CI} = \phi + w^*(\theta).$$

Then, the agent picks $a(\theta) = \phi + w^*(\theta)$ so that for each state θ the same action is induced as in the non-transparent case. Besides, the expected payoff of the designer (over the same states as those considered in the non-transparent case) is $E(\tau)a^{CI} - E(w^*(\theta))a^{CI}$, which is the same as the expected payoff in the non-transparent case, given that it is readily verified that $E(w^*(\theta)) = E(w(\theta))$. Clearly, such identity is non-generic.

states in which the action would be the same in the case in which the agent knows the state.[19]

3) 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 participation constraints are not binding.[20] 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) 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. While dealing with such an environment requires further analysis, I conjecture that the above non-transparency result would still hold in this case.

4) In the context of optimal mechanisms, the most general form of mechanisms allows for the presence of a mediator who could make recommendations to the agent as to which action to choose. When such mechanisms are available, one can always implement the optimal mechanism by having the agent be only informed of what to do (which action a to choose) (see Myerson, 1982). When several states induce the same action as considered in the above Proposition, Myerson's result implies that it is enough to let the agent know what to do and that there is no gain in letting him know more. From that perspective, what the above non-transparency result shows is the stronger property that under the conditions of the above Proposition, the Principal can strictly improve upon what she can achieve in the fully transparent case (which Myerson's result does not imply).

5) Gjesdal (1982) was the first to note that stochastic contracts may sometimes help

[19] Observe that in the derivation of this result, it is not the case that the first effect is made of an order greater than the second (the two effects appear to be of the same magnitude, but for well chosen information structures, the former dominates the latter).

[20] If the participation constraints are binding both at $\theta = \theta_1$ and $\theta_2 + \varepsilon\delta$ in the main argument used to prove the above Proposition when w is set at $w(\theta)$ in state θ, one has to worry that the agent gets no less than his outside option payoff when the agent does not know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$, which may require increasing the burden to the Principal.

the Principal achieve larger expected payoffs. His result though is of a very different nature than the one derived above as it does not rely on variations of the state.[21]

6 On the art of staying silent

The above analysis has illustrated that full transparency is suboptimal whenever the same action would be chosen by the agent for at least two distinct states, which in particular implies that full transparency can be improved upon whenever the state space has a dimensionality greater than the action space.

[21] Roughly, Gjesdal's argument can be illustrated as follows. The reference to the state θ will be omitted from the agent and the principal's utilities, since Gjesdal considers only one state. When the agent chooses a and the principal chooses w, I let $u(a, w)$ denote the utility of the agent and I assume that the payoff of the designer takes the linear form: $\pi(a, w) = a - w$.

That is, the principal wishes to induce a high effort level a from the agent, and she can affect the choice of a using monetary instrument w at some cost assumed to be linear in w.

Gjesdal's main observation is that if $\frac{\partial u}{\partial a}(a, w)$ is convex in w, then a stochastic contract increases the expected utility of the Principal. To see this most simply, assume that instead of using a deterministic w, the principal chooses a stochastic contract that assigns probability $\frac{1}{2}$ to $w_1 = w - \varepsilon$ and probability $\frac{1}{2}$ to $w_2 = w + \varepsilon$. For any given ε, the first-order condition that determines the agent's action is:

$$\frac{1}{2}\frac{\partial u}{\partial a}(a(\varepsilon), w + \varepsilon) + \frac{1}{2}\frac{\partial u}{\partial a}(a(\varepsilon), w - \varepsilon) = 0 \tag{6}$$

Total differentiation of (6) with respect to ε and simple Taylor expansions yield:

$$\frac{da}{d\varepsilon}(\varepsilon) = -\frac{\frac{\partial^3 u}{\partial a \partial w^2}(a(0), w)}{\frac{\partial^2 u}{\partial a^2}(a(0), w)}\varepsilon + o(\varepsilon)$$

where $\frac{o(\varepsilon)}{\varepsilon} \to 0$ as $\varepsilon \to 0$. Thus, $\frac{da}{d\varepsilon}(\varepsilon) > 0$ for ε small enough whenever $\frac{\partial^3 u}{\partial a \partial w^2}(a(0), w) > 0$ (given that $\frac{\partial^2 u}{\partial a^2}(a(0), w) < 0$). This implies that when $\frac{\partial u}{\partial a}(a, w)$ is convex in w, the proposed stochastic contract does strictly better than the original deterministic contract, since the expected monetary transfer is the same in the two cases and the effort level $a(\varepsilon)$ is (slightly) bigger in the stochastic case than in the deterministic case.

Thus, when $\frac{\partial u}{\partial a}(a(0), w)$ is convex in w, any deterministic contract can be improved upon by using a stochastic contract that is a (sufficiently small) mean preserving spread of the original deterministic contract.

Of course, when $\frac{\partial u}{\partial a}(a(0), w)$ is not convex in w, it may be that no stochastic contract dominates the best deterministic contract. So Gjesdal's insight is not that in general optimal contracts are stochastic, but rather than under some convexity conditions they may be stochastic. This is, of course, very different from our non-transparency insight stipulating that when θ has dimension bigger than a, it is always best not to let the agent know perfectly (θ, w).

While characterizing the optimal disclosure policy would be out of reach in general, I wish to investigate the slightly simpler question of when the Principal should say nothing about the state whenever her choice is either to remain silent or to disclose the state θ perfectly to the agent. I note that such a disclosure policy is very easy to implement given that anyone has the option to remain silent.

So the question is: With the aim of maximizing the expected payoff of the Principal, over which set of states should the Principal commit to remaining silent (while disclosing perfectly the state when the latter falls outside this set)?

I will address this question in the context of the leading example introduced in Section 3 in which the state is $\theta = (\beta, \gamma)$, the agent must choose an effort level $a \in \mathbb{R}^+$, the cost of exerting effort a is $c(a) = \frac{1}{2}a^2$, the probability of success is γa, the Principal observes signal $\sigma = 1$ according to $\Pr(\sigma = 1 \mid \textit{success}) = \beta > \frac{1}{2}$ and observes $\sigma = 0$ otherwise, the agent receives a bonus of 1 in case $\sigma = 1$ is observed, and the Principal receives an extra benefit R in case of success.

As already shown in Section 3, conditional on $\theta \in \Theta^*$, the difference in Principal's expected payoff when the agent is uninformed about $\theta \in \Theta^*$ and when he is informed of $\theta \in \Theta$ writes

$$H(\Theta^*) = R[E(\gamma)E(\gamma(2\beta - 1)) - E(\gamma^2(2\beta - 1))] + E(\gamma^2(2\beta - 1)^2) - [E(\gamma(2\beta - 1))]^2$$

where the expectations bear over the realizations of θ conditional on $\theta \in \Theta^*$. Thus the best set of sates Θ^* over which the Principal should remain silent should maximize with respect to Θ^*

$$K(\Theta^*) = \Pr(\theta \in \Theta^*)H(\Theta^*).$$

Considering the marginal effect of the addition (or subtraction) of some small neighborhood of $\theta^* = (\beta^*, \gamma^*)$ to Θ^* yields the following characterization result:

Proposition 4 *The optimal set of states Θ^* over which the Principal should remain silent* *is of the form* $\Theta^* = \Theta^{NE} \cup \Theta^{SW}$ *with*

$$\begin{aligned}\Theta^{NE} &= \{\theta = (\beta, \gamma) \ / \ \gamma - \frac{\gamma(2\beta - 1)}{R} \geq B \textit{ and } \gamma(2\beta - 1) \leq C\} \\ \Theta^{SW} &= \{\theta = (\beta, \gamma) \ / \ \gamma - \frac{\gamma(2\beta - 1)}{R} \leq B \textit{ and } \gamma(2\beta - 1) \geq C\}\end{aligned}$$

where the scalars B *and* C *satisfy:*

$$\begin{aligned}B &= E(\gamma - \frac{\gamma(2\beta - 1)}{R} \mid \theta \in \Theta^*) \\ C &= E(\gamma(2\beta - 1) \mid \theta \in \Theta^*)\end{aligned}$$

The set of states Θ^* over which the Principal should optimally remain silent can be represented graphically. It is depicted in the $(\gamma, \gamma(2\beta - 1))$ space in the dashed area of Figure 1:

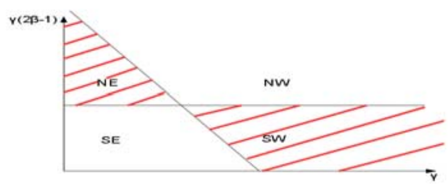

Figure 1

It is good for the Principal to bundle the states in the NE and SW areas because it allows her to shift the relatively high effort level of the agent in the area NE to area SW in which the Principal cares more about effort.[22] Remaining silent also when θ falls in SE would not be a good idea for the Principal because it would lower the overall effort

[22] Effort as measured by $\gamma(2\beta - 1)$ is typically higher in NE than in SW, and effectiveness of effort as measured by γ is typically higher in SW than in NE.

level made in NE and SW, and the Principal cares in aggregate more about effort when $\theta \in NE \cup SW$ than when $\theta \in SE$ given than on average γ is bigger in $NE \cup SW$ than in SE.

7 Conclusion

In this paper, I have shown that some form of non-transparency is desirable, as long as in the complete information benchmark, the same actions would be chosen by the agent(s) for at least two distinct states. Based on this result, I would suggest that some form of non-transparency would seem optimal quite generally, given that in my view the state is likely to vary over more dimensions than the action in most applications. I have also provided for an example the optimal way to remain silent. Clearly, more work should be devoted to understanding the best disclosure policy in more general settings and how it interacts with the best use of monetary instruments.

Appendix

Proof of Proposition 1.

I consider the case in which the action a may have arbitrary many dimensions. Let θ_1 and θ_2 be two states such that in the complete information case $a(\theta_1) = a(\theta_2) = a_1$ and consider a direction δ and ε small enough such that $a(\theta_2 + \varepsilon\delta) \neq a(\theta_2)$.

Let $a_2(\varepsilon)$ and $a^{CI}(\varepsilon)$ be defined by

$$\begin{cases} \nabla_a u(a_1; \theta_1) = 0 \\ \nabla_a u(a_2(\varepsilon); \theta_2 + \varepsilon\delta) = 0 \\ p(\theta_1)\nabla_a u(a^{CI}(\varepsilon); \theta_1) + p(\theta_2 + \varepsilon\delta)\nabla_a u(a^{CI}(\varepsilon); \theta_2 + \varepsilon\delta) = 0 \end{cases} \tag{7}$$

Let $J_u(\theta)$ denote the Jacobian of u at state θ. Total differentiation of (7) yields

$$\begin{aligned} \frac{da_2}{d\varepsilon}(\varepsilon &= 0) = -(J_u(\theta_2))^{-1} \cdot \frac{\partial}{\partial\delta}\nabla_a u(\theta_2) \\ \frac{da^{CI}}{d\varepsilon}(\varepsilon &= 0) = -p(\theta_2)[p(\theta_1)J_u(\theta_1) + p(\theta_2)J_u(\theta_2)]^{-1} \cdot \frac{\partial}{\partial\delta}\nabla_a u(\theta_2) \end{aligned}$$

Let $\Delta(\varepsilon)$ denote the expected organizational gain of not letting the agent know $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ relative to the full information case. We have:

$$\frac{\Delta(\varepsilon)}{\varepsilon} = -p(\theta_1)\nabla_a\pi(a_1;\theta_1)\cdot\frac{da^{CI}}{d\varepsilon}(\varepsilon = 0) + p(\theta_2)\nabla_a\pi(a_1;\theta_2)\cdot[\frac{da_2}{d\varepsilon}(\varepsilon = 0) - \frac{da^{CI}}{d\varepsilon}(\varepsilon = 0)] + o(\varepsilon)$$

where $o(\varepsilon) \to 0$ as $\varepsilon \to 0$.

One can conclude as in the case in which a is one-dimensional by noting that for generic π functions

$$-p(\theta_1)\nabla_a\pi(a_1;\theta_1)\cdot\frac{da^{CI}}{d\varepsilon}(\varepsilon = 0) + p(\theta_2)\nabla_a\pi(a_1;\theta_2)\cdot[\frac{da_2}{d\varepsilon}(\varepsilon = 0) - \frac{da^{CI}}{d\varepsilon}(\varepsilon = 0)] \neq 0 \tag{8}$$

which can be seen by considering the family of Principal's objective

$$\pi_\lambda(a;\theta) = \pi(a;\theta) + \lambda \parallel \theta - \theta_2 \parallel a \cdot \frac{da^{CI}}{d\varepsilon}(\varepsilon = 0)$$

If $\pi(a;\theta)$ does not satisfy (8), $\pi_\lambda(a;\theta)$ will for small non-zero values of λ.

So we may now assume that not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ strictly dominates the full information case. Consider a small η and the effect of not letting the agent know the realization of θ in $B_\eta(\theta_1) \cup B_\eta(\theta_2 + \varepsilon\delta)$ as compared with the full information benchmark where $B_\eta(\theta)$ denotes the set of states which are at (euclidean) distance no more than η from θ. Continuity of π and u implies that conditional on $\theta \in B_\eta(\theta_1) \cup B_\eta(\theta_2 + \varepsilon\delta)$, this difference converges to $\Delta(\varepsilon)/(p(\theta_1) + p(\theta_2 + \varepsilon\delta))$ as η converges to 0 where $\Delta(\varepsilon)$ is the function described above. It follows that there must exist η small enough so that the Principal is strictly better off not letting the agent know $\theta \in B_\eta(\theta_1) \cup B_\eta(\theta_2 + \varepsilon\delta)$ as compared with the full information benchmark. **Q. E. D.**

Proof of Proposition 2.

To present the argument in a more reader-friendly way, I restrict myself to the case in which the action space has dimension 1 and there are two agents i and j.

For any θ, $a^{NE}(\theta)$ must solve:

$$\begin{cases} \frac{\partial u_i}{\partial a_i}(a_i, a_j; \theta) = 0 \\ \frac{\partial u_j}{\partial a_j}(a_i, a_j; \theta) = 0 \end{cases}$$

Consider θ_1 and θ_2 such that $a_i^{NE}(\theta_1) = a_i^{NE}(\theta_2)$ and consider a direction δ of θ such that for ε small enough $a^{NE}(\theta_2 + \varepsilon\delta) \neq a^{NE}(\theta_1)$.

If agent i does not know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$, NE actions $a_i^C(\varepsilon)$, $a_{j,1}^C(\varepsilon)$ and $a_{j,2}^C(\varepsilon)$ are given by:

$$\begin{cases} \frac{\partial u_j}{\partial a_j}(a_i^c(\varepsilon), a_{j,1}^c(\varepsilon); \theta_1) = 0 \\ \frac{\partial u_j}{\partial a_j}(a_i^c(\varepsilon), a_{j,2}^c(\varepsilon); \theta_2 + \varepsilon\delta) = 0 \\ p(\theta_1)\frac{\partial u_i}{\partial a_i}(a_i^c(\varepsilon), a_{j,1}^c(\varepsilon); \theta_1) + p(\theta_2 + \varepsilon\delta)\frac{\partial u_i}{\partial a_i}(a_i^c(\varepsilon), a_{j,2}^c(\varepsilon); \theta_2 + \varepsilon\delta) = 0 \end{cases}$$

And if there is full information, NE actions $a_{i,1}$, $a_{i,2}(\varepsilon)$, $a_{j,1}$ and $a_{j,2}(\varepsilon)$ are given by:

$$\begin{cases} \frac{\partial u_j}{\partial a_j}(a_{i,1}, a_{j,1}; \theta_1) = 0 \\ \frac{\partial u_j}{\partial a_j}(a_{i,2}(\varepsilon), a_{j,2}(\varepsilon); \theta_2 + \varepsilon\delta) = 0 \\ \frac{\partial u_i}{\partial a_i}(a_{i,1}, a_{j,1}; \theta_1) = 0 \\ \frac{\partial u_i}{\partial a_i}(a_{i,2}(\varepsilon), a_{j,2}(\varepsilon); \theta_2 + \varepsilon\delta) = 0 \end{cases}$$

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

$$\begin{aligned} \Delta(\varepsilon) &= p(\theta_1)[\pi(a_i^c(\varepsilon), a_{j,1}^c(\varepsilon); \theta_1) - \pi(a_{i,1}, a_{j,1}; \theta_1)] + \\ &\quad p(\theta_2 + \varepsilon\delta)[\pi(a_i^c(\varepsilon), a_{j,2}^c(\varepsilon); \theta_2 + \varepsilon\delta)] - \pi(a_{i,2}(\varepsilon), a_{j,2}(\varepsilon); \theta_2 + \varepsilon\delta)] \end{aligned}$$

Similarly to the one agent case if $\Delta'(0) \neq 0$, then it implies that not letting agent i know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ 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 $\pi_\lambda(a_i, a_j; \theta) = \pi(a_i, a_j; \theta) + \lambda a_i \parallel \theta - \theta_1 \parallel^2$ where $\parallel \cdot \parallel$ denotes the Euclidean distance.

Proof of Proposition 3.

To make the argument more reader-friendly, assume the action of the agent a varies over one dimension and let w_σ denote the bonus obtained by the agent when the Principal observes σ ex post. For any θ, there is an optimal $w = (w_\sigma)_\sigma$, say $w(\theta)$. It is defined as

$$\begin{aligned} w(\theta) &= \arg\max_w \pi(a(w), w; \theta) \\ a(w) &= \arg\max_a u(a, w; \theta) \end{aligned}$$

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

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

Define $\widetilde{\pi}(a;\theta) = \pi(a, w(\theta);\theta)$ and $\widetilde{u}(a;\theta) = u(a, w(\theta);\theta)$. Apply the argument developed around Proposition 1 assuming $\widetilde{\pi}$ is the designer's objective and $\widetilde{u}$ is the agent's utility function. Clearly, if not letting the agent know whether $\theta = \theta_1$ or $\theta_2 + \varepsilon\delta$ 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 $w(\theta)$ in state θ).

It remains to show that generically it is not the case that (see (4) above)

$$\frac{\partial\widetilde{\pi}}{\partial a}(\theta_1)\frac{\partial^2\widetilde{u}}{\partial a^2}(\theta_2) - \frac{\partial\widetilde{\pi}}{\partial a}(\theta_2)\frac{\partial^2\widetilde{u}}{\partial a^2}(\theta_1) = 0 \tag{9}$$

To see this, consider the family of π_λ functions

$$\pi_\lambda(a, w;\theta) = \pi(a, w;\theta) + \lambda \parallel \theta - \theta_1 \parallel^2 (a + \sum_\sigma \frac{\partial^2 u/\partial a\partial w_\sigma}{\partial^2 u/\partial a^2}(a_2, w(\theta_2);\theta_2)\cdot w_\sigma)$$

where $\lambda \in \mathbb{R}$. For such a family, $w(\theta)$ are the same at $\theta = \theta_1$ (resp. θ_2) whatever λ so that $\frac{\partial\widetilde{\pi}_\lambda}{\partial a}(\theta) = \frac{\partial\widetilde{\pi}}{\partial a} + \lambda \parallel \theta - \theta_1 \parallel^2$ for $\theta = \theta_1$ and θ_2. Thus, if $\widetilde{\pi}$ satisfies (9), $\widetilde{\pi}_\lambda$ does not for any sufficiently small $\lambda \neq 0$, and one can conclude as in the case of Proposition 1. **Q. E. D.**

Proof of Proposition 4

The proof takes the following route. Consider a candidate set Θ^* over which the Principal is supposed to optimally stay silent. Let $p(\Theta^*)$ be the probability that θ falls in Θ^*. Let $\theta^* = (\beta^*, \gamma^*)$ be a state in Θ and assume the probability that $\theta = \theta^*$ is p^* (I will later on explain how one deals with non-atomic distributions p).

Assume first that θ^* lies outside Θ^*. Letting $K(S)$ be the expected organizational gain over full transparency when the Principal remains silent over states $\theta \in S$, routine calculations show that $K(\Theta^* \cup \{\theta^*\}) - K(\Theta^*)$ divided by $p(\Theta^*)p^*R$ writes:

$$-[\gamma^* - E(\gamma \mid \Theta^*) - \frac{a(\theta^*) - E(a(\theta) \mid \Theta^*)}{R}]\cdot[a(\theta^*) - E(a(\theta) \mid \Theta^*)] \tag{10}$$

where $a(\theta) = \gamma(2\beta - 1)$.

Clearly if (10) is strictly positive, the Principal is strictly better off being silent over $\Theta^* \cup \{\theta^*\}$ than over Θ^* and thus Θ^* is not optimal (in the case of smooth densities, replacing $\{\theta^*\}$ by a small neighborhood of θ^* would do using continuity arguments). This implies that for all $\theta^* \notin \Theta^*$, one should have:

$$[\gamma^* - E(\gamma \mid \Theta^*) - \frac{a(\theta^*) - E(a(\theta) \mid \Theta^*)}{R}] \cdot [a(\theta^*) - E(a(\theta) \mid \Theta^*)] \geq 0$$

Assume next that $\theta^* \in \Theta^*$. Then the same calculation yields that $K(\Theta^*) - K(\Theta^* \setminus \{\theta^*\})$ has the same sign as

$$-[\gamma^* - E(\gamma \mid \Theta^*) - \frac{a(\theta^*) - E(a(\theta) \mid \Theta^*)}{R}] \cdot [a(\theta^*) - E(a(\theta) \mid \Theta^*)]$$

and thus the optimality of Θ^* implies that for every $\theta^* \in \Theta^*$, one should have:

$$[\gamma^* - E(\gamma \mid \Theta^*) - \frac{a(\theta^*) - E(a(\theta) \mid \Theta^*)}{R}] \cdot [a(\theta^*) - E(a(\theta) \mid \Theta^*)] \leq 0$$

It follows that the boundary of the optimal silence set Θ^* is of the form depicted in Proposition 4 in which the boundary between Θ^* and $\Theta \setminus \Theta^*$ is defined by

$$[\gamma^* - E(\gamma \mid \Theta^*) - \frac{a(\theta^*) - E(a(\theta) \mid \Theta^*)}{R}] \cdot [a(\theta^*) - E(a(\theta) \mid \Theta^*)] = 0$$

Letting

$$\begin{aligned} B &= E(\gamma \mid \Theta^*) + \frac{E(a(\theta) \mid \Theta^*)}{R} \\ C &= E(a(\theta) \mid \Theta^*) \end{aligned}$$

and replacing $a(\theta)$ by $\gamma(2\beta - 1)$ yields the desired conclusion. **Q. E. D.**

References

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

[2] Eso, P and B. Szentes (2007): 'Optimal information disclosure in auctions,' *Review of Economic Studies* **74**, 705-731.

[3] Gjesdal, F. (1982): 'Information and Incentives: the Agency Information Problem,' *Review of Economic Studies* **49**, 373-390.

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

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

[6] 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.

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

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

[9] Milgrom, P. and R. Weber (1982): 'A theory of auctions and competitive bidding,' *Econometrica*, **50**, 1089-1122.

[10] Milnor, J. W. (1965): *Topology from the differentiable viewpoint* Princeton University Press.

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

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

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

[14] Tamura, W. (2012): 'A theory of multidimensional information disclosure,' mimeo.