OPTIMAL REGULATION OF COOPERATIVE R&D UNDER INCOMPLETE INFORMATION[*]

Isabelle Brocas†

A regulator offers a cooperation contract to two firms to develop a research project. The contract provides incentives to encourage skill-sharing and coordinate subsequent efforts. Innovators must get informational rents to disclose their privately known skills, which results in distorting R&D efforts with respect to the first-best level. When efforts are strategic complements, both efforts are distorted downwards. By contrast, when efforts are strategic substitutes, the effort of the firm with most valuable skills is distorted downwards (to decrease rents) and the effort of the other firm is distorted upwards (to compensate the previous efficiency loss).

I. INTRODUCTION

THE OUTCOME OF A RESEARCH AND DEVELOPMENT (R&D) ACTIVITY is generally affected by the amount of resources allocated (effort, time, money, etc.) and a series of intrinsic characteristics measuring the efficiency of innovative firms. Those features can be of very different nature: ability (skills of workers), expertise, efficiency of internal routines, previous knowledge acquired through transfers of technologies, licenses or patents, etc. However, they all have the property of increasing the probability of success of an R&D project.

In a context of R&D rivalry, the amount of resources is usually not socially optimal and the transfer of technologies is not ensured. To mitigate these inefficiencies, policymakers have developed a series of tools directed to encourage and promote cooperative research. Privately organized research partnerships, like Research Joint Ventures, receive government funding and joint research is sponsored by public authorities (some examples are MCC in the United States and MITI in Japan). In Europe and Japan, subsidized cooperative R&D projects have become an important tool of the industrial policy, with programs like ESPRIT, BRITE, BIOTECH, EUREKA, RACE, VLSI, etc. The aim is to improve the economic and social impact of R&D activities by ensuring a better dissemination of results and the transfer of technologies from various sources.

Despite the potential benefits of cooperation, the impossibility of observing the skills of each partner, of ensuring the transfer of knowledge, and of monitoring the allocation of resources impairs the performance of cooperative agreements from a social viewpoint. In this paper, we focus our attention on the design of optimal contractual rules selected by a benevolent regulator (referred to as she) whose aims are to favor efficient sharing of skills between invited firms and to encourage socially optimal provision of effort, in a context of incomplete information. More precisely, we build on regulation theory (see, e.g., Baron and Myerson [1982] and Laffont and Tirole [1986]) and consider a model in which a welfare maximizing regulator offers a cooperation contract to two firms in order to develop a research project. The contract consists of a system of transfers that includes taxes and subsidies raised through distortionary taxation. The difficulty for the design of an optimal contract relies on the fact that, while disclosure-contingent transfers are required to encourage the sharing of skills among team members, they must not distort the incentives for adequate development efforts.

Cooperative arrangements have been extensively studied in the literature. For the purpose of this paper, we will classify previous works in three categories. In the first one (the skill-sharing literature), the analyses concentrate on skill-sharing. Skills are defined very broadly. They include patents, expertise of research personnel, previous R&D results, organizational routines, etc. Skill-sharing refers to a transfer of skills between firms that improves the R&D outcome. The definition is relevant for firms with homogeneous skills (they share common areas of expertise, even though they might use different but comparable techniques) heterogeneous skills (they have different areas of expertise that need to be used in conjunction to achieve the project) and both homogeneous and heterogeneous skills (they share some but not all the areas of expertise). The theoretical and empirical analyses show that the possibility of developing synergies from the exchange of technical knowledge is an important motive for firms to make cooperative agreements (Hagedoorn [1993]).[1]

In the second category (the effort-coordination literature), authors have mostly focused on situations where firms are symmetric in terms of their skills, which is the case if they come from the same industry, and are equally capable of embark in R&D. The papers highlight the idea that cooperation allows firms to share their costs, realize economies of scale and avoid over- and under-provision of effort in the industry (Katz [1986], d'Aspremont and

---

[1] Hamel [1991] highlights the fact that cooperative arrangements are opportunities for each partner to internalize the skills of others and create next generation competencies. This knowledge is not easily diffused across firms and cooperative agreements are necessary to facilitate learning. These agreements also allow firms to benefit from R&D inputs that are not tradable, like routines (see Arrow [1962]), or for which markets are imperfect, like personnel. Besides, cooperation allows a quick learning of skills that would otherwise require time (Milgrom and Roberts [1990]). See also Sakakibara [1997] for an empirical analysis.

Jacquemin [1988] and Katz and Ordover [1990]). Overall, those analyses concentrate on the effect of cost-sharing, and show that it corrects the market failures that prevent firms from investing optimally in R&D. However, the papers do not generally treat the issue of asymmetric information.

In the third category (the optimal contracting literature), team members design contracts under asymmetric information (skills are not observable) and/or moral hazard (costs cannot be monitored). Bhattacharya, Glazer and Sappington [1990,1992] study optimal skill-sharing when team members have perfectly substitutable skills and embark separately on R&D. Given the competitive framework in the R&D stage, firms with high expertise may be reluctant to share it with their counterparts, because doing so puts all the team members on equal footing in the research stage. The authors analyze the intra-team transfers and licensing mechanism that ensure an efficient sharing of skills. D'Aspremont, Bhattacharya and Gérard-Varet [1998] also study optimal intra-team payments. Contrary to the other works, firms not only cooperate at the skill-sharing level but they also embark jointly in R&D. The paper derives conditions under which the first-best outcome for the team can be implemented despite the existence of incomplete information and moral hazard.[2]

This paper incorporates aspects of the three previously mentioned categories. First, building on the skill-sharing literature, our paper is the first theoretical work that compares optimal skill-sharing contracts when firms have homogeneous and perfectly transferable skills, homogeneous but imperfectly transferable skills and heterogenous (complementary) skills. Second, as in the effort-coordination literature, we quantify the efforts that are generated in a cooperative agreement when firms embark jointly on R&D. Third, as in the contracting literature, we determine the optimal contract given the combination of moral hazard, adverse selection, and interdependent payoffs between agents. Unlike in previous works (with the exception of d'Aspremont et al. [1998] and Gandal and Scotchmer [1993]), our contract combines revelation and sharing of private knowledge with coordination of efforts between team members. Furthermore, the crucial difference between our paper and all the previous literature (including the two papers above mentioned) is that our objective is not to determine which mechanism implements the first-best outcome (effort and skill-sharing) from the team's perspective given costless intra-team transfers. Instead, we analyze the contract that maximizes social welfare given that transfers from

---

[2] The paper uses the methodology developed by d'Aspremont and Gérard-Varet [1979] and Arrow [1979] to study implementation of efficient decisions for the production of public goods under incomplete information. Team problems have also been studied in Gandal and Scotchmer [1993]. The authors analyze the incentives to implement the joint-profit maximizing solution with and without budget balance when firms' abilities are private information. All these works build on the seminal paper on moral hazard in teams by Holmström [1982].

taxpayers to firms are socially costly. Due to asymmetry of information and costly transfers, the contract implemented is necessarily a second-best optimum, in which the regulator solves a trade-off between rents and efficiency. Our contribution is then to characterize the optimal second-best efforts and compare them with the complete information case.

We now proceed to a brief overview of our main results. In sections 3 and 4, we show that disclosure of skills is costly, so the regulator must grant informational rents to the innovators in exchange of their valuable information. This increases the marginal social cost of the innovation and creates an incentive to distort R&D efforts with respect to the first-best outcome, the standard trade-off between rent and efficiency (Propositions 1 and 2). More precisely and other things being equal, the regulator wants to decrease the effort of each innovator. However, the *equilibrium* pair of efforts will depend on the degree of strategic substitutability of efforts. When efforts are strategic complements, a decrease in the effort of one innovator decreases the marginal benefit of effort of its partner. This, combined to the increased marginal cost due to informational rents, implies that the second-best equilibrium effort of both firms is smaller than in the first-best case (Proposition 3), as in the standard one-agent literature. By contrast, when efforts are strategic substitutes, decreasing the effort of one innovator increases its partner's marginal benefit of effort. This encourages the effort of the latter which, contrary to all the existing literature, might end up being distorted upwards in equilibrium (Proposition 4). In other words, the regulator may optimally decrease the effort of one agent relative to the first-best level in order to reduce the informational rents and, to compensate for this efficiency loss, increase the effort of the other agent. This occurs only if the strategic substitutability of efforts is sufficiently strong and if the marginal cost of informational rents is sufficiently small.

In section 5 we investigate in more detail the cases in which we are likely to observe an upward distortion of effort, since this is the main novelty of the paper. Interestingly, we show that such distortion may occur both under homogeneous and heterogeneous skills, and that the rationale is the same in both cases. Basically, the cost in terms of informational rents of inducing a firm to disclose its private information is increasing in the value of its skills. In order to reduce rents, the regulator then substitutes effort of the firm with most valuable skills with effort of the firm with least valuable skills. For the homogeneous case, this means an effort below and above the first-best level for the innovator with highest and lowest (ex-post revealed) skills respectively (Proposition 5). For the heterogeneous case, it means an effort below and above the first-best level for the innovator whose skill is most and least essential in the sharing function respectively (Proposition 6).

The implications of our model are simple yet, in our opinion, powerful. First, and contrary to usual practices by policy-makers like the European Commission, transfers that promote cooperative ventures are effective only

if they establish specific splitting rules of the payments among the team members. Second, it is crucial that the regulator determine which team member has the most valuable skills (either because its input is more needed or because its knowledge is more developed) in order to encourage skill-sharing. This can be achieved with an appropriate system of transfers (e.g., higher payoffs for members who prove their higher relative value). Third, the regulator must demand more effort from the least valuable members than from the most valuable ones to compensate for their lower skills. This and other implications are discussed in the following sections.

The paper is organized as follows. Section II presents the model. The optimal contracts offered to innovators under complete and incomplete information are characterized in sections III and IV respectively. Section V analyzes in detail the effect of the nature of skills on the optimal outcome. Last, section VI concludes.

II. THE MODEL

II(i). *Basic Ingredients*

We consider two risk-neutral firms or innovators 1 and 2, indexed by i and j. The efficiency of innovator i to complete research projects is represented by the parameter $\theta_i \in [\underline{\theta}_i, \overline{\theta}_i]$, where $\underline{\theta}_i$ and $\overline{\theta}_i$ are respectively the smallest and highest possible levels. For the time being and for simplicity, we refer to θ_i as the 'skills' or more generically the 'type' of innovator i, and we assume that it reflects the ability or degree of expertise of employees, the efficiency of internal routines, the knowledge, etc. In section II(iii), we extensively discuss different interpretations of this parameter and how each interpretation affects the specific modelling of the game. Innovators have private information about their own type θ_i. The type of firm i is drawn, independently of the type of firm j, from a common knowledge distribution $F_i(\theta_i)$ strictly increasing, continuously differentiable, with density $f_i(\theta_i)$ and such that $F_i(\underline{\theta}_i) = 0$ and $F_i(\overline{\theta}_i) = 1$. To avoid bunching phenomena, we assume (as usual in the contracting literature) that $F_i(\cdot)$ satisfies the standard monotone hazard rate property.

Assumption 1. $\frac{1-F_i(\theta_i)}{f_i(\theta_i)}$ is decreasing in θ_i for all i.

The R&D cooperation game is divided in three stages.

Stage 1. Contract.

The regulator (she) wants to finance a cooperative research project for which innovators 1 and 2 have to share their skills. She does not observe the types of innovators and offers the following take-it-or-leave-it contract. Innovators

are asked to report their types, and the regulator commits to give a transfer
to each of them contingent on the reports and the observed outcome of the
research activity.[3] The contract also includes punishments if one innovator
accepts the contract but fails to share its skills, provided it can be observed or
verified by the regulator (this assumption is discussed in Remark 1 below). If
one firm refuses the contract, the game ends and there is no research activity.
If they both accept, firms make reports $\tilde{\theta}_1 \in [\underline{\theta}_1, \overline{\theta}_1]$ and $\tilde{\theta}_2 \in [\underline{\theta}_2, \overline{\theta}_2]$ and the
contract is implemented. The reservation utility of each innovator is
normalized to zero.

Stage 2. Sharing of skills.

We consider situations where the private information parameter of at least
one firm is relevant to the other firm's research activity. Since the contract is
observable by all parties, if innovator i reports a type $\tilde{\theta}_i$ in stage 1, then
innovator j expects to share $\tilde{\theta}_i$. There are two cases to consider. If $\tilde{\theta}_1 \theta_1$ and
$\tilde{\theta}_2 \theta_2$, innovators share $\tilde{\theta}_1$ and $\tilde{\theta}_2$ and proceed to stage 3. If at least one
innovator over-represents its type ($\tilde{\theta}_i > \theta_i$), it cannot share it. In that case,
the project is delayed. We assume that, should this occur, the regulator
observes this delay, shuts down the project and inflicts an arbitrarily large
penalty on the firm. As a consequence, whenever innovator i plans to report
$\tilde{\theta}_i > \theta_i$, it has to make sure that it will be able to share $\tilde{\theta}_i$ in stage 2, that is, it
must be able to 'upgrade its skills' to $\tilde{\theta}_i$. The cost of acquiring $\tilde{\theta}_i$ for a firm of
type $\theta_i (< \tilde{\theta}_i)$ is denoted $c(\theta_i, \tilde{\theta}_i)$ and satisfies the following assumption.

Assumption 2. $c(\theta_i, \theta_i) = 0$ and $\partial c(\theta_i, \tilde{\theta}_i)/\partial \tilde{\theta}_i \geqslant 0$ and $\partial c(\theta_i, \tilde{\theta}_i)/\partial \theta_i \leqslant 0$ for
all $\tilde{\theta}_i > \theta_i$.

Note that Assumption 2 embraces as a particular case the situation in which
it is impossible to overstate the type (for this, we simply need to set $c(\tilde{\theta}_i, \theta_i) = +\infty$ for all $\tilde{\theta}_i > \theta_i$).

Overall, the skills of firm i at the end of stage 2 is a combination of its own
final type $\tau_i = max(\tilde{\theta}_i, \theta_i)$ and the type reported and shared by the other firm
$\tilde{\theta}_j$. We denote it by $m_i(\tau_i, \tilde{\theta}_j)$. Cooperation is formalized by assuming that
firms benefit from sharing their skills.

Assumption 3. $\partial m_i(\tau_i, \tilde{\theta}_j)/\partial \tau_i \geqslant 0$, $\partial m_i(\tau_i, \tilde{\theta}_j)/\partial \tilde{\theta}_j \geqslant 0$ and $m_i(\tau_i, \tilde{\theta}_j) \geqslant \tau_i$
for all $\tau_i, \tilde{\theta}_j$.

---

[3] More specifically, we analyze direct revelation mechanisms. In practice, the resulting
optimal contract can be implemented with a wide variety of (direct and indirect) mechanisms,
depending on the characteristics of the project at hand. From a theoretical perspective and
using the revelation principle, we know that the principal cannot obtain a greater payoff than
with the optimal direct mechanism. Our analysis thus provides a normative benchmark.

This assumption captures the idea that each innovator benefits from having high skills and also from its partner having high skills, as long as these are shared. In section II(iii) below, we discuss how different interpretations of the parameter θ_i lead to different representations of the sharing function $m_i(\cdot, \cdot)$.

Stage 3. Research and Development.

Each innovator implements a non-observable R&D effort $e_i (\geqslant 0)$. Denote by $e = (e_1, e_2)$ the vector of efforts chosen by firms. Innovator i suffers a disutility $\psi(e_i, m_i)$ of exerting effort e_i, where $\psi(\cdot, \cdot)$ is common knowledge and satisfies the following conditions.

Assumption 4. $\psi(e_i, m_i)$ is increasing and convex in e_i and decreasing in m_i. Moreover, $\psi(0, m_i) = 0$ for all m_i and $\psi_{12}(e_i, m_i) < 0$ for all e_i and m_i.

The assumption that skills decrease the cost of effort is standard in contract theory (see Remark 2 below).[4] Note furthermore that the cost of each incremental unit of effort is decreasing in firm i's final skills m_i ($\psi_{12} < 0$).[5]

Efforts affect the result $\gamma (\in \{S, F\})$ of the research stage. More precisely, we assume that with probability $\pi(e_1, e_2)$ the cooperative project is successful ($\gamma = S$), in which case an innovation of fixed and known social value V is obtained (where V is sufficiently big so that it is socially desirable to encourage some effort). With probability $1 - \pi(e_1, e_2)$, the project fails ($\gamma = F$) and no innovation is obtained. This probability satisfies the following conditions.

Assumption 5. $\pi(e_1, e_2)$ is increasing, concave in e_1 and e_2 and symmetric in its arguments (i.e., $\pi(e^a, e^b) = \pi(e^b, e^a)$).[6]

This assumption reflects the idea that sharing efforts induces synergies. For instance, the probability of success when firm 1 receives no help from firm 2 is $\pi(e_1, 0)$. Under cooperation, success occurs with probability $\pi(e_1, e_2) > \pi(e_1, 0)$. Moreover, the effort of each firm affects symmetrically the probability of success ($\pi(e^a, e^b) = \pi(e^b, e^a)$). This implies in particular that there is no exogenous reason for the regulator to encourage more effort from one innovator rather than the other. This assumption helps us to

---

[4] See for instance Bhattacharya, Glazer and Sappington [1992] and d'Aspremont, Bhattacharya and Gérard-Varet [1998].

[5] As usual in moral hazard contexts, 'effort' and 'disutility of effort' must be interpreted in a broad sense. That is, it may represent the amount of time and resources spent by the firm to develop a project, the number of employees allocated to perform a given task, etc. Hence, a firm that employs more talented workers (higher θ) needs to allocate fewer of them in the project (lower e) to obtain the same output.

[6] This means in particular that $\pi_1(e^a, e^b) = \pi_2(e^b, e^a)$ and $\pi_{11}(e^a, e^b) = \pi_{22}(e^b, e^a)$.
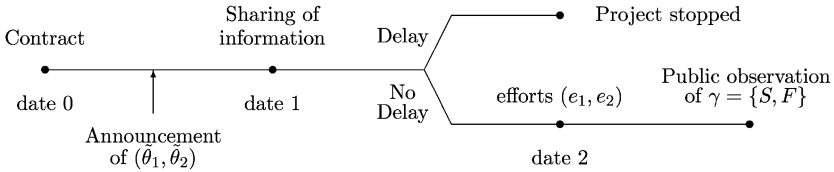
Figure 1
Timing

isolate the effects of skill-sharing on optimal efforts. Last, the result γ of the research stage is publicly observed and innovators receive their payments. The overall timing of the game is summarized in Figure 1.

Remark 1: Over-representation of type. The standard contracting literature rules out by assumption the possibility that firms over-state their type when they participate in cooperative projects (see e.g., Bhattacharya, Glazer and Sappington [1990,1992] and d'Aspremont, Bhattacharya and Gérard-Varet [1998]). In our view, the capacity to over-state depends crucially on the ability of the regulator to punish firms, which is linked to the nature of the type. If it consists of verifiable private information, such as a patent or a license, over-estimation is prevented by fixing an appropriate punishment whenever it is observed.[7] By contrast, if the private information is not costlessly verifiable (e.g., claiming that an employee is the best expert in his field), then the regulator has to include a credible punishment to avoid over-estimation, which itself depends on the difficulty of verifying the information. In this model, we are assuming a zero or relatively small cost of verification (the regulator does not observe if over-representation occurs but observes its consequences), so that threatening firms with shutting down the project is sufficient. Such punishments are frequently used in practice. For example, the European Commission verifies if the timetable is respected and threatens firms with cutting their financing if outcomes are a long time coming. Alternatively, when it is difficult to check over-reporting, the regulator needs to resort to more sophisticated monetary incentives, such as penalties and rewards for denouncing the partner. In any case, these procedures only prevent deviations, and are not used in equilibrium. Therefore, although the degree of verifiability and the types of punishments to use are important issues in practice, our results do not rely on them. Last, recall that the traditional assumption that firms can only under-report their type reduces to $c(\tilde{\theta}_i, \theta_i) = +\infty$ for all $\tilde{\theta}_i > \theta_i$ in our model.

Remark 2: Modelling skills. In stage 3, we assume that higher skills decrease the innovator's cost of effort and does not affect its probability of success. It

[7] Given perfect verifiability, over-representation is observed with probability one.

would be formally equivalent to assume that skills increase the probability of success without affecting the cost of effort. Our choice is made purely for tractability reasons.[8]

Remark 3: Complementary or substitutability of efforts. It is important to realize that Assumption 5 only requires synergies between the efforts of firms, i.e. a probability of success increasing and weakly concave in both arguments. It thus embraces two conceptually different cases. First, the case of effort substitutability, where the marginal effect of innovator i's effort on the probability of success decreases as the effort of innovator j increases (or, formally, $\pi_{12}(e_1, e_2) < 0$). Second, the case of effort complementarity, where the marginal effect of innovator i's effort on the probability of success increases as the effort of innovator j increases (or, formally, $\pi_{12}(e_1, e_2) > 0$).[9]

II(ii). *The Revelation Mechanism*

The regulator and innovators play the following game. Innovators simultaneously announce their parameter $\tilde{\theta}_i$, which determines the type to be shared. We assume that innovators never observe the true type of their partner (although, in the optimal contract offered by the regulator, each innovator will in equilibrium truthfully reveal it). For each vector of reports $\tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2)$, the contract offered by the regulator specifies for each firm i a transfer $t_i^S(\tilde{\theta})$ if the project succeeds ($\gamma = S$), and a transfer $t_i^F(\tilde{\theta})$ if the project fails ($\gamma = F$). Last, conditional on innovators reporting their true type, i is expected to choose an effort level $\hat{e}_i(\tilde{\theta})$.

   In this sequential game, the strategy of innovator i is a pair of functions $(\sigma_i, \tilde{e}_i)$. First, i chooses a report that depends on its true type $\tilde{\theta}_i \equiv \sigma_i(\theta_i)$. Second, i exerts an effort which is a function of its final type τ_i and the type announced by both innovators $\tilde{\theta}$, that is $\tilde{e}_i(\tilde{\theta}, \tau_i)$. Those strategies must form an equilibrium. We look for a perfect Bayesian Nash equilibrium. As in the standard contracting literature, it is defined as follows.

Definition 1. The strategies $(\sigma_1, \sigma_2, \tilde{e}_1, \tilde{e}_2)$ form a perfect Bayesian equilibrium if:

 (i) The announcement functions (σ_1, σ_2) form a Bayesian equilibrium in the game defined by transfers $(t_1^S, t_1^F, t_2^S, t_2^F)$, efforts $(\tilde{e}_1, \tilde{e}_2)$ and the prior beliefs characterized by $F_1(\cdot)$ and $F_2(\cdot)$;

---

[8] Again, this assumption is standard in the literature. See for instance d'Aspremont, Bhatacharya and Gérard-Varet [1998].

[9] For example, suppose that $(e_1, e_2) \in [0, 1]^2$. If $\pi(e_1, e_2) = (e_1 + e_2 - e_1 e_2)/3$ then $\pi_{12} < 0$ and efforts are substitutes. If $\pi(e_1, e_2) = (e_1 + e_2 + e_1 e_2)/3$ then $\pi_{12} > 0$ and efforts are complements.

(ii) For all vector of announcements $\tilde{\theta}$, efforts $(\tilde{e}_1, \tilde{e}_2)$ form a Bayesian equilibrium of the subgame in which innovator i's updated belief over j's true type is $\tilde{F}_j(\theta_j | \tilde{\theta}_j)$;

(iii) $\tilde{F}_j(\cdot \,|\, \tilde{\theta}_j)$ is such that if $\sigma_j(\theta_j) \neq \tilde{\theta}_j$, the probability that j's type is θ_j is zero.

Given the equilibrium strategies of firms, the regulator's revelation contract is optimal if innovator i is induced to disclose its true type θ_i and to select the expected effort level $\hat{e}_i(\theta)$.

Definition 2. The compensations $(t_1^S(\tilde{\theta}), t_1^F(\tilde{\theta}), t_2^S(\tilde{\theta}), t_2^F(\tilde{\theta}))$ allow perfect disclosure and the selection of the expected efforts $(\hat{e}_1(\theta), \hat{e}_2(\theta))$ if:

(i) There exists a Bayesian equilibrium $(\sigma_1, \sigma_2, \tilde{e}_1, \tilde{e}_2)$ such that $\sigma_i(\theta_i) = \theta_i$ for all i;

(ii) Let $(\sigma_1, \sigma_2, \tilde{e}_1, \tilde{e}_2)$ be a Bayesian equilibrium such that $\sigma_i(\theta_i) = \theta_i$ for all i. Then, the transfers $(t_1^S(\theta), t_1^F(\theta), t_2^S(\theta), t_2^F(\theta))$ lead to the selection of $\tilde{e}_i(\theta, \theta_i) \equiv \tilde{e}_i(\theta) = \hat{e}_i(\theta)$ for all i.

Given Definitions 1 and 2, we can proceed to the analysis of the optimal contract offered by the regulator in this model of cooperation. However, we would like first to discuss possible interpretations of the firm's type and alternative models of the skill-sharing function.

II(iii).   *Interpretations and Models of 'Skills'*

Our general point is that, by sharing their type θ (efficiency, ability, information, expertise, knowledge, etc.), firms benefit from the 'skills' of their partner (Assumption 3) which improves the outcome of the R&D activity (Assumptions 4 and 5 combined). Then θ represents the 'capital of skills' of the firm in each of its area(s) of expertise. In practice, however, each example requires a different model for the sharing function $m_i(\cdot, \cdot)$.[10] To be more precise, consider two firms cooperating in a project. Each firm has several areas of expertise, some of them overlap and some of them do not, some of them can be combined and some of them are mutually exclusive. Then, sharing skills can imply two conceptually different things. First, it may account for the capacity of firms to improve their performance in the areas of expertise that are common (e.g., by sharing their knowledge and their researchers). Second, it may refer to the adoption of new valuable expertise in the other areas.[11] Interestingly, both effects are present in the thematic research projects organized by the European Commission (ESPRIT, BRITE,

[10] I am grateful to the editor and an anonymous referee for pointing out this issue and suggesting the discussion.

[11] Firms may also be concerned with the future use by the other firm of the acquired expertise. Our model does not study this issue.

BIOTECH, EUREKA and RACE). These programs aim at stimulating technological innovation, encouraging traditional sectors of industry to incorporate new processes and developing scientific and technological collaboration within and across different areas of science and technology. For instance, BIOTECH focuses on 8 research areas related to life technologies (e.g., genome analysis, plant and animal biotechnology) and the European Commission provides funding to projects in the context of the programme. It stimulates cooperation between experts in the same domain, but also across fields.

To isolate and compare the two effects, we analyze two extreme situations. In the first one, there is one area of expertise and one relevant skill needed to complete the project. It therefore corresponds to the case of homogeneous and substitutable skills. The parameter θ_i represents the level of skills of firm i in this specific area. In the second one, there are several areas of expertise and multiple types of skills that can be combined to complete the project. It therefore corresponds to the case of heterogeneous and complementary skills. The parameter θ_i represents the level or skills of firm i in its own area of expertise.

The first situation is particularly relevant for cooperative projects developed between firms in the same industry. One of the firms usually has a more advanced technology, better machinery, more up-to-date-knowledge, etc. (which is what we generically label as type or skill). It therefore transmits this know-how to the other firm. Examples of single industry cooperation are VLSI (Very Large Scale Integrated Circuit) developed in Japan between 1975 and 1985, and its American counterpart SEMATECH (1987). Both projects were consortia of semiconductor manufacturers and were designed to help manufacturers catch up with semiconductor technology. SEMATECH was originally created to re-invigorate the U.S. semiconductor industry, and has evolved into the world's premiere research consortium. Member companies cooperate precompetitively in key areas of semiconductor technology. Their common aim is to accelerate development of the advanced manufacturing technologies that will be needed to build tomorrow's most powerful semiconductors. Homogeneity of skills is captured by assuming that types can be ranked and only the lowest-type firm benefits from sharing. If types are perfect substitutes (i.e., they can be ordered in the sense of Blackwell) then we have $m_i(\tau_i, \tilde{\theta}_j) = max\{\tau_i, \tilde{\theta}_j\}$. In our model, we consider a broader class of functions to account for a positive but limited capacity of firms to assimilate new technologies. This is summarized in the next assumption and depicted in Figure 2a (note that $\alpha = 1$ corresponds to the case where types are perfectly transferable).[12]

$m_1(\tau_1, \tilde{\theta}_2)$ $m_1(\tau_1, \tilde{\theta}_2)$

a $\underline{\theta}_1$ $\tilde{\theta}_2$ τ_1 b $\underline{\theta}_1$ τ_1
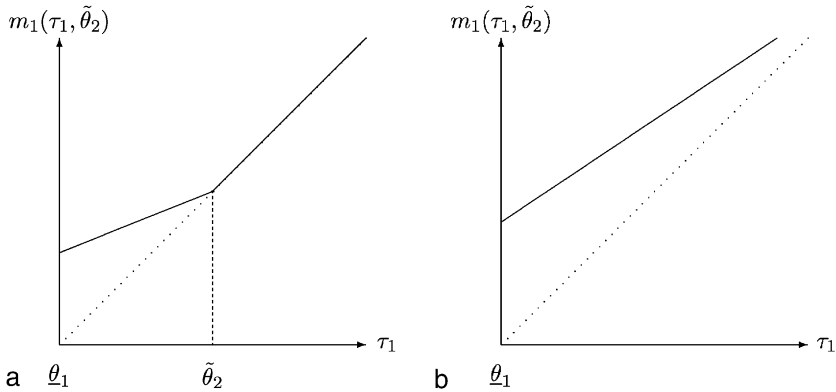
Figure 2
a Homogeneous types, b Heterogeneous types

Assumption 6. When skills are homogeneous, then $m_i(\tau_i, \tilde{\theta}_j) = (1 - \alpha)\tau_i + \alpha\tilde{\theta}_j$ if $\tau_i < \tilde{\theta}_j$ and $m_i(\tau_i, \tilde{\theta}_j) = \tau_i$ if $\tau_i \geqslant \tilde{\theta}_j$, where $\alpha \in (0, 1]$.

The second situation is more relevant in projects where the success depends on the combination of expertise in different and complementary areas of science and technology. For example, fiber optics communication systems were developed by combining optics and electronics. Probably the best-known case of a project requiring cooperation in different fields is the ISS (International Space Station), the largest scientific cooperative program in history, that draws on the resources and expertise of 16 nations. For instance, the project creates unique cross-disciplinary research programs, bringing the basic sciences of physics, biology, and chemistry together with a wide range of engineering disciplines. It aims at understanding living systems, by conducting experiments to study biological and chemical processes that cannot be conducted on earth.

Other innovations depend on research in a unique area but combine qualitatively different types of skills. This is the case when one innovator is specialized in fundamental research (a public research center, a university) while the other has acquired expertise in development (industry). In this vein, European thematic projects involve firms, laboratories and universities. In all these cases, skills are heterogeneous and their combination makes the strength of the venture. A simple way of capturing this effect is proposed in the next assumption and illustrated in Figure 2b.

Assumption 7. When skills are heterogeneous, then $m_1(\tau_1, \tilde{\theta}_2) = \beta\tau_1 + \tilde{\theta}_2$ and $m_2(\tau_2, \theta_1) = \beta\tilde{\theta}_1 + \tau_2$, where $\beta(> 1)$ represents the weight associated to innovator 1.

In the next sections we will study the optimal cooperation contract under both homogeneity and heterogeneity of skills.

III. COOPERATION UNDER COMPLETE INFORMATION

We begin the analysis by characterizing the optimal contract offered by a benevolent regulator in the benchmark case of complete information, i.e., when she can observe the types of innovators and monitor their efforts. The regulator maximizes social welfare, which in our case, is simply the sum of the utilities of innovators and consumers. Following the standard regulation theory (see e.g., Laffont and Tirole [1986]), we assume that transfers are raised through distortionary taxes. Each unit of money received by an innovator costs society $1 + \lambda$, where $\lambda > 0$ represents the shadow cost of public funds.[13]

If the regulator offers transfers $t_i^S(\theta)$ and $t_i^F(\theta)$ to innovator i in case of success and failure respectively, its expected utility is:

$$u_i(\theta) = t_i^S(\theta)\pi(e_1, e_2) + t_i^F(\theta)(1 - \pi(e_1, e_2)) - \psi(e_i, m_i(\theta))$$

Given λ and the social value of the project V, the expected surplus of consumers Π is:

$$\begin{aligned} \Pi = {} & \pi(e_1, e_2)V - (1 + \lambda)[\pi(e_1, e_2)[t_1^S(\theta) + t_2^S(\theta)] \\ & + (1 - \pi(e_1, e_2))[t_1^F(\theta) + t_2^F(\theta)]] \end{aligned}$$

Last, the social welfare W is the sum of the consumers surplus and the utilities of both firms:

$$\begin{aligned} W = {} & \Pi + u_1(\theta_1, \theta_2) + u_2(\theta_1, \theta_2) \\ = {} & \pi(e_1, e_2)V - (1 + \lambda)[\psi(e_1, m_1(\theta_1, \theta_2)) + \psi(e_2, m_2(\theta_1, \theta_2))] \\ & - \lambda[u_1(\theta_1, \theta_2) + u_2(\theta_1, \theta_2)] \end{aligned}$$

From the definition of W we note that, if λ is positive, the ex-post utility of firms enters negatively in the social welfare function. Also, recall that firms accept the regulator's contract if and only if $u_i(\theta_1, \theta_2) \geqslant 0$. Therefore, under complete information, the regulator will adjust the transfers under success and failure t_i^S and t_i^F so as to make sure that $u_i(\theta_1, \theta_2) = 0$. The next (sufficient) condition guarantees the existence of a unique equilibrium in the R&D stage.

Assumption 8. $\pi_{11}(e_1, e_2) < -|\pi_{12}(e_1, e_2)|$. Moreover, the third derivatives of the functions $\pi(\cdot, \cdot)$ and $\psi(\cdot, \cdot)$ are zero.

Our first result is a characterization of optimal efforts under complete information.

---

[13] The parameter λ reflects the costs of satisfying the budget constraint of the government. When $\lambda = 0$, the budget constraint of the government is not binding. In that case, it is not socially costly to finance public projects.

Proposition 1. Under complete information, firm i's socially optimal effort $e_i^*(\theta)$ is such that:

$$\pi_i(e_1^*(\theta), e_2^*(\theta)) V - (1 + \lambda) \psi_1(e_i^*(\theta), m_i(\theta)) = 0$$

where $\frac{\partial e_1^*}{\partial \theta_1} > 0$, $\frac{\partial e_2^*}{\partial \theta_2} > 0$, $\frac{\partial e_1^*}{\partial \lambda} < 0$ and $\frac{\partial e_2^*}{\partial \lambda} < 0$.

Proof: See Appendix 1.

Under complete information, the regulator induces the innovators to exert the first-best level of effort. This effort is such that its marginal social benefit is equal to its marginal social cost. Naturally, the higher the type of an innovator, the cheaper the cost of exerting effort. Then, the regulator induces each innovator to exert an effort increasing in its type. Besides, since transfers are socially costly ($\lambda > 0$), the optimal efforts are decreasing in the shadow cost of public funds.

Naturally, innovators must be compensated by the regulator for their effort. Recall that they only accept contracts such that $u_i(\theta) \geqslant 0$ and that the utility of innovators enters negatively the social welfare function. Thus, the optimal transfers $t_i^{S*}(\theta)$ and $t_i^{F*}(\theta)$ to firm i in case of success and failure respectively satisfy the following equality:

$$t_i^{S*}(\theta)\pi(e_1^*(\theta), e_2^*(\theta)) + t_i^{F*}(\theta)(1 - \pi(e_1^*(\theta), e_2^*(\theta))) - \psi(e_i^*(\theta), m_i(\theta)) = 0$$

Before characterizing the optimal (second-best) contract under incomplete information, let us make the following remark. For all θ, denote by $e_i^{FB}(e_j, \theta)$ innovator's i effort reaction function under complete information, that is its first-best effort as a function of the effort of the partner. This function satisfies the following first order condition:

$$\pi_i(e_i^{FB}(e_j, \theta), e_j)V - (1 + \lambda)\psi_1(e_i^{FB}(e_j, \theta), m_i(\theta)) = 0$$

Differentiating with respect to e_j, we get:

$$\frac{de_i^{FB}}{de_j} = -\frac{\pi_{12}(e_i^{FB}(e_j, \theta), e_j)V}{\pi_{ii}(e_i^{FB}(e_j, \theta), e_j)V - (1 + \lambda)\psi_{11}(e_i^{FB}(e_2, \theta), m_i(\theta))}$$

Note that $\pi_{12}(e_1, e_2) \gtrless 0 \Leftrightarrow \frac{de_i^{FB}}{de_j} \gtrless 0$. In words, when $\pi_{12}(e_1, e_2) < 0$ the optimal effort of firm i is decreasing in the effort of firm j so efforts are strategic substitutes. When $\pi_{12}(e_1, e_2) > 0$ the optimal effort of firm i is increasing in the effort of firm j so efforts are strategic complements.

IV.   COOPERATION UNDER INCOMPLETE INFORMATION

The analysis becomes more interesting when we assume that the regulator cannot observe the types of innovators and cannot monitor their efforts. In section IV(i), we determine the properties that the incentive scheme must

satisfy in order to induce innovators to reveal truthfully and share efficiently their type. In section IV(ii), we characterize the optimal contract offered by the regulator to the innovators under incomplete information and compare it with the full information contract presented in section IV(iii). In section IV(iv), we provide some comments on the modelling and suggest some extensions.

IV(i).   *Incomplete Information and Incentives*

We look for a Bayesian Nash equilibrium for the direct mechanism where innovators are asked to report their private information parameters (see Definitions 1 and 2). This means that for every vector θ, firm i prefers to report honestly its type if the partner also does. We denote by $\Phi_i(\tilde{\theta}_i, \theta_i)$ the expected utility of firm i at the beginning of the game when its type is θ_i and it reports $\tilde{\theta}_i$:

$$
\begin{aligned}
\Phi_i(\tilde{\theta}_i, \theta_i) = E_{\theta_j}[& t_i^S(\tilde{\theta}_i, \theta_j)\pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tau_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j)) \\
& + t_i^F(\tilde{\theta}_i, \theta_j)[1 - \pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tau_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))] \\
& - \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tau_i), m_i(\tau_i, \theta_j))] - c(\theta_i, \tilde{\theta}_i)\mathbb{1}_{\tau_i = \tilde{\theta}_i}
\end{aligned}
$$

where $\mathbb{1}_{\tau_i = \tilde{\theta}_i} = 1$ if $\tau_i = \tilde{\theta}_i$ and $\mathbb{1}_{\tau_i = \tilde{\theta}_i} = 0$ otherwise. The expected utility of firm i when both innovators report truthfully is:

$$
U_i(\theta_i) = \Phi_i(\theta_i, \theta_i)
$$

Also, we denote by $u_i(\theta)$ the utility of firm i in the third stage when both firms report truthfully their type:

$$
\begin{aligned}
u_i(\theta) = E_{\theta_j}[& \pi(\tilde{e}_i(\theta, \tau_i), \tilde{e}_j(\theta, \tau_j))t_i^S(\theta) \\
& + [1 - \pi(\tilde{e}_i(\theta, \tau_i), \tilde{e}_j(\theta, \tau_j))]t_i^F(\theta) - \psi(\tilde{e}_i(\theta, \tau_i), m_i(\tau_i, \theta_j))]
\end{aligned}
$$

Given asymmetric information and moral hazard, the optimal revelation contract must satisfy a number of constraints. First, to induce truth-telling, the contract must be such that any deviation from disclosure provides a smaller utility to the firm (incentive compatibility):

(IC) $\Phi_i(\theta_i, \theta_i) \geqslant \Phi_i(\tilde{\theta}_i, \theta_i)$

Second, the regulator cannot force firms to accept the contract. The expected utility of each innovator must be greater than its reservation utility (ex-ante individual rationality).

(IR$_{ea}$) $U_i(\theta_i) \geqslant 0$

Third, the contract must be such that the expected utility of each firm in stage 3 is also greater than the reservation utility (ex-post individual rationality).

(IR$_{ep}$) $u_i(\theta_i, \theta_j) \geqslant 0$

Fourth, the regulator cannot monitor the efforts implemented by the firms. The contract must be such that firm i chooses the expected effort level (moral hazard).

(MH) $\quad \tilde{e}_i(\theta) = \underset{e_i}{\mathrm{argmax}}\ \pi(e_i, \tilde{e}_j(\theta)) t_i^S(\theta) + (1 - \pi(e_i, \tilde{e}_j(\theta))) t_i^F(\theta)$

$$- \psi(e_i, m_i(\theta))$$

Overall, the regulator's optimization program $\mathcal{P}$ under asymmetric information and moral hazard is:

$$\mathcal{P}: \underset{\{t_i^S(\theta), t_i^F(\theta), \tilde{e}_i(\theta)\}}{max} \qquad W$$

$$\text{s.\,t.} \qquad (\text{IC}) - (\text{IR}_{\text{ea}}) - (\text{IR}_{\text{ep}}) - (\text{MH})$$

The following assumption guarantees that the cost of acquiring type $\tilde{\theta}_i > \theta_i$ is greater than its benefits in terms of disutility saving. It allows us to ensure that the conditions for truthful revelation are sufficient.[14]

Assumption 9. There exists $k > 0$ such that $|\psi_2(e, m)| < k$ for all e and m. Moreover for all $\tilde{\theta}_i > \theta_i$, $\partial c(\theta_i, \tilde{\theta}_i)/\partial \theta_i < -k \cdot \partial m_i(\tilde{\theta}_i, \theta_j)/\partial \tilde{\theta}_i$.[15]

Following the usual techniques, we get the first preliminary result.

Lemma 1. The regulator's optimization program $\mathcal{P}$ can be rewritten as:

$$\underset{\{\tilde{e}_1, \tilde{e}_2\}}{max}\ E_\theta[\pi(\tilde{e}_1(\theta), \tilde{e}_2(\theta)) V - (1 + \lambda)(\psi(e_1(\theta), m_1(\theta))$$

$$+ \psi(e_2(\theta), m_2(\theta)))] - \lambda(E_{\theta_1} U_1(\theta_1) + E_{\theta_2} U_2(\theta_2))$$

(IC_1) $\quad$ s.t. $\qquad \dfrac{dU_i}{d\theta_i}(\theta_i) = -E_{\theta_j}\left[\psi_2(\tilde{e}_i(\theta), m(\theta))\dfrac{dm_i}{d\theta_i}\right]$

(IC_2) $\quad \dfrac{\partial \tilde{e}_i}{\partial \theta_i}(\theta) \geqslant 0$

(IR_{ep}) $\quad u_i(\theta_i, \theta_j) \geqslant 0$

Proof: See Appendix 2.

These are the standard conditions in mechanism design problems. In the R&D (third) stage, innovator i selects the effort that maximizes its expected

---

[14] The idea is simple. If the cost of acquiring $\tilde{\theta}_i$ is compensated by the gains of using a more efficient skill during the research stage, then the regulator wants to induce firms to acquire it. Since we want to concentrate on a standard model of information revelation, we prevent this from occurring. Overall, in this model we assume implicitly that firms have already exploited all the opportunities to acquire skills so that acquiring new skills is sufficiently costly (higher skills do not exist, are not tradable or require time, research and money to be obtained). For an analysis of optimal contracting when the regulator aims at promoting research to acquire skills before embarking on an R&D project, see Brocas [2002].

[15] Note that k can be arbitrarily large.

utility, as given by (MH). In the contract (first) stage, it makes a report anticipating the effort level it will exert later. To induce truthful revelation (that is, to satisfy (IC)), the regulator must give informational rents to the innovators. Note that an innovator with a type θ_i can always pretend to have a type $\theta_i < \theta_i$ and save on effort. In order to prevent such behavior, both the effort and the informational rents of each innovator must be increasing in its type. This is reflected in (IC$_1$) and (IC$_2$), where $dU_i/d\theta_i \geqslant 0$ and $\partial \tilde{e}_i /\partial \theta_i \geqslant 0$. Last, note that the innovator accepts the contract in the first stage only if it does not have incentives to leave in the second stage. Then, (IR$_{ea}$) is automatically satisfied if (IR$_{ep}$) holds.

From (IC$_1$), we deduce that firm i's equilibrium expected utility is:

$$U_i(\theta_i) = -\int_{\underline{\theta}_i}^{\theta_i} \int_{\underline{\theta}_j}^{\overline{\theta}_j} \psi_2(\tilde{e}_i(s,\theta_j), m_i(s,\theta_j))\frac{\partial m_i}{\partial s} dF_j(\theta_j)ds + U_i(\underline{\theta}_i)$$

Since the utility of firms enters negatively the social welfare function (W is decreasing in $U_i(\theta_i)$), the regulator will minimize the rents. Note that $U_i(\theta_i) = \int_{\underline{\theta}_j}^{\overline{\theta}_j} u_i(\theta_i, \theta_j)dF_j(\theta_j)$. Therefore, by setting $U_i(\underline{\theta}_i) = 0$ rents are minimized and the constraint (IR$_{ep}$) is satisfied:

$$u_i(\theta_i, \theta_j) = -\int_{\underline{\theta}_i}^{\theta_i} \psi_2(\tilde{e}_i(s,\theta_j), m_i(s,\theta_j))\frac{\partial m_i}{\partial s} ds \geqslant 0$$

Replacing the equilibrium expressions of $U_i(\theta_i)$ in the objective function of the regulator and integrating by parts, we can rewrite the social welfare function in the following way:

$$\begin{aligned} W' = E_{\theta_1}E_{\theta_2}\Big[& \pi(\tilde{e}_1(\theta), \tilde{e}_2(\theta))V - (1+\lambda)(\psi(e_1(\theta), m_1(\theta)) \\ & + \psi(e_2(\theta), m_2(\theta))) + \lambda\Big(\psi_2(e_1(\theta), m_1(\theta))\frac{1-F_1(\theta_1)}{f_1(\theta_1)}\frac{dm_1}{d\theta_1} \\ & + \psi_2(e_2(\theta), m_2(\theta))\frac{1-F_2(\theta_2)}{f_2(\theta_2)}\frac{dm_2}{d\theta_2}\Big)\Big] \end{aligned}$$

The last term in the function W' reflects the fact that the regulator must grant a rent to innovator i in order to induce disclosure of its private information parameter.

IV(ii).   *The Optimal Contract*

We are now in a position to determine the optimal mechanism. From Lemma 1 and the subsequent analysis, we know that the optimal contract maximizes W' under the remaining constraint (IC$_2$). Optimal efforts are characterized in the next proposition.

Proposition 2. Under incomplete information, firm i's optimal effort $\hat{e}_i(\theta)$ is such that:

$$\pi_i(\hat{e}_1(\theta), \hat{e}_2(\theta))V - (1 + \lambda)\psi_1(\hat{e}_i(\theta), m_i(\theta))$$
$$= -\lambda\psi_{12}(\hat{e}_i(\theta), m_i(\theta))\frac{1 - F_i(\theta_i)}{f_i(\theta_i)}\frac{dm_i}{d\theta_i}$$

where $\frac{\partial \hat{e}_1}{\partial \theta_1} > 0$ and $\frac{\partial \hat{e}_2}{\partial \theta_2} > 0$.

Proof: See Appendix 3.

Proposition 2 reflects a trade-off between efficiency and rent extraction. If the regulator were concerned exclusively with efficiency, she would compare the social benefit $\pi(e_1, e_2)V$ and cost $(1 + \lambda)\psi(e_i, m_i)$ of efforts, just like in Proposition 1. However, under incomplete information, she also needs to grant an informational rent $-\lambda\psi_2(e_i, m_i)\frac{1-F_i(\theta_i)}{f_i(\theta_i)}\frac{dm_i}{d\theta_i}$ to the innovator in order to induce the desired effort level. Given that the marginal cost of effort decreases with the firm's type ($\psi_{12} < 0$), these rents are increasing in the effort demanded to the innovator. The second-best contract trades-off optimal efforts and rent minimization.

Denote by $R(e_i)$ the marginal cost of granting informational rents to innovator i and by $e_i^{SB}(e_j, \theta)$ the effort reaction function of innovator i under asymmetric information (i.e., its second-best effort as a function of the partner's effort). Formally:

$$R(e_i) = -\lambda\psi_{12}(\hat{e}_i(\theta), m_i(\theta))\frac{1 - F_i(\theta_i)}{f_i(\theta_i)}\frac{dm_i}{d\theta_i} (\geqslant 0)$$

$$\pi_i(e_i^{SB}(e_j, \theta), e_j)V - (1 + \lambda)\psi_1(e_i^{SB}(e_j, \theta), m_i(\theta)) = R(e_i^{SB}(e_j, \theta))$$

From these definitions, we immediately observe that $R(e_i) = 0$ if $\theta_i = \overline{\theta}_i$ and $R(e_i) > 0$ if $\theta_i < \overline{\theta}_i$. This implies that $e_i^{SB}(e_j, \overline{\theta}_i, \theta_j) = e_i^{FB}(e_j, \overline{\theta}_i, \theta_j)$ and $e_i^{SB}(e_j, \theta_i, \theta_j) < e_i^{FB}(e_j, \theta_i, \theta_j)$ for all $\theta_i < \overline{\theta}_i$. In words, under asymmetric information, the highest-type firm exerts the first-best effort level (the traditional 'no distortion at the top' result). By contrast, holding the partner's effort fixed, the effort of all the other types is distorted downwards: efficiency is decreased in order to diminish the rents. Last, note that the willingness of the regulator to distort efforts when $\theta_i < \overline{\theta}_i$ together with Assumption 1 implies that the effort increases in the type, as required by (IC$_2$).

The second-best efforts described above can be implemented in the optimal contract. Basically, the regulator must ensure first that the expected utility given the transfers is high enough so that the innovators accept the contract, and second that the difference between the transfer in case of success and the transfer in case of failure is such that innovators select the desired effort levels $\hat{e}_i(\theta)$. These two conditions are satisfied by the transfers $\{\hat{t}_i^S(\theta), \hat{t}_i^F(\theta)\}$,

which are fully characterized in Appendix 3. In particular, the transfer received in case of success is greater than the transfer received in case of failure.[16]

We have thus determined the equilibrium efforts under incomplete information and shown that the regulator can design a contract and specify transfers that implement these efforts. We also know that the effort reaction function of each innovator $e_i^{SB}(\cdot)$ is below the first-best reaction function $e_i^{FB}(\cdot)$. However, this does not necessarily imply that the *equilibrium* efforts will always be smaller than under complete information. The next two results deal precisely with this issue.

Proposition 3. When efforts are strategic complements ($\pi_{12} > 0$), optimal efforts under incomplete information are such that $\hat{e}_1(\theta)e_1^*(\theta)$ and $\hat{e}_2(\theta)e_2^*(\theta)$ for all θ.

Proof: See Appendix 4.

The intuition is simple. As mentioned in Proposition 2, under incomplete information, each innovator's marginal cost of effort is increased by the presence of informational rents. Other things equal, the regulator wants to decrease the effort of firms with respect to the first-best solution, to minimize rents. Moreover, given strategic complementarity, a decrease in the effort of one innovator results in a decrease in its partner's marginal benefit of effort. It therefore provides a further incentive to reduce effort. Overall, solving the trade-off between rents and efficiency for the two innovators simultaneously results in a downward distortion of both efforts. In equilibrium, $\hat{e}_1(\theta) \leqslant e_1^*(\theta)$ and $\hat{e}_2(\theta) \leqslant e_2^*(\theta)$, as depicted in Figure 3.

When efforts are strategic substitutes, the optimal efforts under incomplete information have the following properties.

Proposition 4. When efforts are strategic substitutes ($\pi_{12} < 0$), optimal efforts under complete information are such that either $\hat{e}_i(\theta) \leqslant e_i^*(\theta)$ for all i or $\hat{e}_i(\theta) \leqslant e_i^*(\theta)$ and $\hat{e}_j(\theta) \geqslant e_j^*(\theta)$.

Proof: See Appendix 4.

As usual, for each effort level of innovator i, the regulator has incentives to decrease the effort of innovator j relative to its first-best level in order to reduce its informational rents $R(e_j)$. Given strategic substitutability, diminishing e_j increases innovator i's marginal benefit of effort. When the strategic substitutability is sufficiently strong, the increase in the marginal benefit of innovator i's effort offsets the cost of the informational rents left to

---

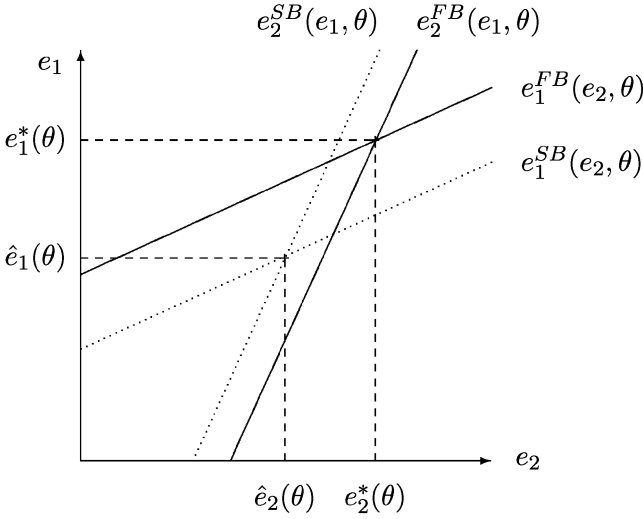[16] Otherwise, firms would maximize the likelihood of failure.

Figure 3
Optimal efforts when $\pi_{12} > 0$

i due to asymmetric information $R(e_i)$. As a result, the regulator induces an effort above the first-best level for that innovator and below the first-best level for the other one. Stated differently, when the regulator decreases the effort of an innovator to reduce the informational rents, she wants to compensate the loss in efficiency by increasing the effort of the partner. This incentive is increasing in the degree of the strategic substitutability of efforts.

In equilibrium, we will observe that either one effort is distorted downwards and the other upwards (when the effort substitution effect is strong) or both efforts are distorted downwards (when the effort substitution effect is weak). Naturally, the regulator will never distort both efforts upwards. The two cases are illustrated in Figure 4a and 4b respectively.

The main message of this section is that the specific modeling of the cooperation game will crucially affect the comparison between efforts under complete and incomplete information. When efforts are strategic complements, it is unambiguously beneficial for the regulator to solve the trade-off efficiency vs. rents with a downward distortion of both efforts. More interestingly, when efforts are strategic substitutes, one of the two efforts may be distorted upwards. The degree of substitutability $\pi_{12}(\cdot, \cdot)$ and the marginal cost of informational rents $R(\cdot)$ determine whether this occurs in equilibrium. Since informational rents depend on the sharing function $m(\cdot, \cdot)$, the likelihood of observing an upward distortion will also depend on several other factors. First, the nature of the parameter θ_i (homogeneous vs. heterogeneous skills, see section II(iii)). Second, the relative levels of the innovators' types (high θ vs. low θ). And third, the distribution of types in the
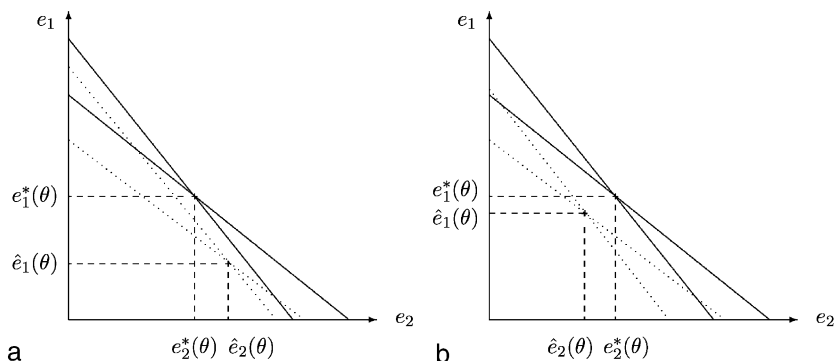
Figure 4
Optimal efforts when $\pi_{12} < 0$

economy (more weight on high types vs. more weight on low types). As the reader can notice, in Proposition 4 we have remained deliberately vague in our analysis of the forces pushing towards a double downward distortion or a two-side distortion. In section V we provide a thorough discussion of how each of the factors mentioned above affects the equilibrium efforts as well as the practical implications for our examples of cooperative projects.

IV(iii). *Remarks and Extensions*

Before closing this section, we should like to address some additional issues.[17] From a positive perspective, our model describes a cooperative arrangement between a regulator and two firms to develop a project entirely financed by the former. Moreover, given the assumptions, the model is best interpreted as a public cooperative project that cannot be achieved in the absence of intervention, either because the individual budget constraints are binding or because the innovation is socially valuable but does not generate high private profits.[18] We now briefly mention what happens when we consider alternative settings.

1. Partially financed contracts. The analysis extends to partially financed programs. To see this, suppose that firms can use the innovation to manufacture products or sell licenses. Formally, we assume that innovators

---

[17] This section builds on the comments and suggestions of the editor and two anonymous referees, to whom I am grateful.

[18] Some fields are traditionally supervised by public authorities which delegate research projects and provide funding. This is true for space research (e.g., ISS), but also for several areas of biology. In this last case, it can be motivated by the necessity to prevent some industrial applications. For instance, the HUMAN GENOME PROJECT aims at conducting research but also at addressing the ethical, legal and social issues that may arise from the project, before transferring technologies to the private sector.

1 and 2 capture a share ρ_1 and ρ_2 of the social surplus, with $1 - \rho_1 - \rho_2 (\geqslant 0)$ being the share captured by consumers. The expected utility of innovator i is then:

$$u_i(\theta) = (t_i^S(\theta) + \rho_i V)\pi(e_1, e_2) + t_i^F(\theta)(1 - \pi(e_1, e_2))$$

$$- \psi(e_i, m_i(\theta))$$

and the social welfare is:

$$W = \pi(e_1, e_2)(1 + \lambda\rho_1 + \lambda\rho_2)V - (1 + \lambda)[\psi(e_1, m_1(\theta))$$

$$+ \psi(e_2, m_2(\theta))] - \lambda[u_1(\theta) + u_2(\theta)]$$

This model can be interpreted as a grand contract between two R&D institutions and an agency that aims at determining the appropriate provision of incentives for research (i.e., optimal efforts) given the share ρ_i of social surplus captured by firms.[19] Just as before, a system of transfers $\{\tilde{t}_i^S(\theta), \tilde{t}_i^F(\theta)\}$ is selected by the agency to implement the optimal outcome.

Overall, the formal model is technically very similar to the previous one. The only difference is that, since firms capture directly a fraction of the benefits of the innovation, the regulator can now achieve specific effort levels with smaller payments. Equivalently, if the regulator is willing to invest a given amount in the project, she can now command higher effort levels. In other words, the principal substitutes transfers that are costly due to the distortionary properties of taxation ($\lambda > 0$) with a share of surplus which is costless because it is directly captured by firms. This means in particular that the direct transfer in case of success can now be smaller than in case of failure ($\tilde{t}_i^S(\theta) \gtrless \tilde{t}_i^F(\theta)$).[20] This also means that the optimal (first and second best) efforts are higher.[21]

2. Privately organized R&D. In our model, an innovator who refuses the contract cannot independently decide to undertake the project (i.e., its outside option is fixed). One could assume instead that firms also have the possibility of proceeding on their own. In that case, their outside option would be increasing in their type or, in other words, the minimum utility that the regulator would be forced to grant them in order to induce acceptance of the contract would be greater the higher their private information parameter. This alternative modelling would modify the shape of the

---

[19] Replacing V by $(1 + \lambda\rho_1 + \lambda\rho_2)V$, we get qualitatively the same results.

[20] Naturally, the firm's total payoff in case of success is always greater than in case of failure: $\tilde{t}_i^S(\theta) + \rho_i V > \tilde{t}_i^F(\theta)$.

[21] It is easy to show that when $\lambda = 0$, i.e., when raising funds through taxation is costless (the budget constraint of the government is always satisfied), the optimal first best efforts are the same in the two cases. Besides, the first best efforts are implemented under asymmetric information and moral hazard and the transfers are $\tilde{t}_i^S(\theta) = \hat{t}_i^S(\theta) - \rho_i V$ and $\tilde{t}_i^F(\theta) = \hat{t}_i^F(\theta)$.

equilibrium transfers. However, the qualitative properties of the effects highlighted in sections III and IV(ii) would still hold.[22]

3. Cooperation. The paper does not address the issue of duplication of effort, defined as spending unnecessary resources for the development of an already existing knowledge. Instead, it focuses on how to efficiently coordinate research. Thus, in this paper, cooperation refers to two specific mechanisms. First, the optimal coordination of efforts as in Katz (1986), d'Aspremont and Jacquemin (1988), Katz and Ordover (1990) and all the effort-coordination literature reviewed in the introduction. Second, the maximal sharing of skills as in Bhattacharya, Glazer and Sappington (1990,1992).

4. Implementation. Our model considers a direct revelation mechanism (the regulator asks innovator *i* to report its type θ_i). In practice, the regulator finances cooperative projects on the basis of reported expertise and reported cost. The cost associated to the project is represented by the disutility function in our model. This means that innovator *i* is asked to disclose its efficiency θ_i as well as its cost $C_i(\theta) = \psi(\tilde{e}_i(\theta), m_i(\theta))$. The mechanism is then formally equivalent to ours. However, in some cases, the regulator is forced to finance a project on the basis of the observed cost only. Does this create an implementation problem? The answer is, not necessarily. Suppose for example that $C = \psi(e, m) = \psi(e - m)$. In that case, $e - m = \psi^{-1}(C)$ and therefore observing the cost C is sufficient to induce the effort e given θ. As long as costs are simple 'functions' of effort and type, the optimal contract can be implemented with transfers that are contingent exclusively on the reported cost.[23]

5. Individual payments vs. team payments. In our setting, the regulator offers a payment to each participant. An alternative modelling of this game would be to assume that the regulator rewards the activity of the team with a transfer $T^S(\tilde{\theta})$ in case of success and a transfer $T^F(\tilde{\theta})$ in case of failure. Then, the two partners design an intra-team contract describing how this transfer is split among them. When designing her contract, the regulator would then take into account the anticipated splitting rule of the team. Imposing this additional constraint would lead to an additional distortion, and therefore a suboptimal contract from the welfare perspective relative to the asymmetric information case described in Proposition 2 (third-best solution).

---

[22] Contracting with type-dependent reservation utilities is an interesting (and technically complex) theoretical issue (for an in-depth analysis, see e.g., Jullien [2000]). The main effect of relaxing the fixed outside option assumption is that the individual rationality constraint does not necessarily bind at the bottom (i.e., at $\theta_i = \underline{\theta}_i$), so informational rents are not necessarily monotonic on the agent's type.

[23] It is usually assumed in regulation theory that the observable cost is linear in effort and type. Considering a more general class of function generally requires additional assumptions to implement the optimal mechanism with a scheme based only on reported costs (see Laffont and Tirole [1990]).

V.   EFFICIENCY AND THE NATURE OF SKILLS

In this section, we analyze in more detail the equilibrium efforts under asymmetric information as a function of the nature of skills, the private information of firms, and the distribution of types. As we have shown in section IV(ii), the effort of one of the innovators can be distorted upwards only if efforts are strategic substitutes. We will thus work under that assumption.

Assumption 10. $\pi_{12}(e_1, e_2) < 0$.

In order to better compare equilibrium efforts, we assume that the type of both innovators is drawn from the same distribution with support $[\underline{\theta}, \overline{\theta}]$. Also, since we will analyze how the shape of the distribution affects the efforts, we need to introduce the following definition.

Definition 3. Distribution $G(\cdot)$ is less favorable than distribution $F(\cdot)$ on $[\underline{\theta}, \overline{\theta}]$ if $G = M(F)$ with M increasing and concave (this implies in particular that $G(\theta_i) \geqslant F(\theta_i)$ for all θ_i).

We analyze separately the case of homogeneous and heterogeneous skills.

V(i).   *Homogeneous Skills*

Suppose that there is one main skill needed to complete the project, types can be ranked and the sharing function satisfies Assumption 6. If, after truthful revelation, it turns out that $\theta_i > \theta_j$, then the skill-sharing functions are $m_i(\theta) = \theta_i$ and $m_j(\theta) = \alpha\theta_i + (1 - \alpha)\theta_j$. The optimal efforts are as follows.

Proposition 5. When skills are homogeneous and given $\theta_i > \theta_j$, there exists $\overline{\alpha}(< 1)$ such that:

(i) If $\alpha > \overline{\alpha}$, then $\hat{e}_j > e_j^*$ and $\hat{e}_i < e_i^*$.
(ii) If $\alpha < \overline{\alpha}$, then $\hat{e}_j < e_j^*$. Moreover, $\hat{e}_i > e_i^*$ if and only if $\theta_i \gg \theta_j$.

Last, if $\alpha = 1$, then $\hat{e}_j^F > \hat{e}_j^G > e_j^*$ and $\hat{e}_i^F < \hat{e}_i^G < e_i^*$.[24]

Proof: See Appendix 5.

When $\alpha = 0$, innovators cannot share their skills, so no benefits are generated at stage 2. Since equilibrium efforts are increasing in the firm's private information parameter ($\partial\hat{e}_i/\partial\theta_i > 0$), the effort of the least efficient innovator will be distorted more heavily than the effort of the most efficient

---

[24] Superscripts F and G in $\hat{e}$ are used to denote that θ_i is drawn from distributions F and G, respectively.

one. This means that, in equilibrium, either both firms exert an effort below the first-best level or the lowest-type firm exerts an effort below and the highest-type firm above the first-best level ($\hat{e}_j(\theta) < e_j^*(\theta)$ and $\hat{e}_i(\theta) \gtreqless e_i^*(\theta)$ when $\theta_i > \theta_j$). This last possibility occurs only if θ_i is sufficiently large compared to θ_j. However, the effect is not very surprising as it relies exclusively on the strategic substitutability of efforts.

More interestingly, when $\alpha = 1$, sharing skills induces strong synergies. In this case, only the information of the most efficient firm is relevant in the R&D stage ($m_i = \max\{\theta_1, \theta_2\}$). Then, an innovator receives rents for the disclosure of its private information only if, ex-post, its type is greater than that of its partner. As a consequence, solving the trade-off between rents and efficiency amounts to distorting only the effort of the most efficient firm.[25] Formally, if ex-post $\theta_1 > \theta_2$, the optimal efforts induced by the regulator are:

$$\pi_1(\hat{e}_1(\theta), \hat{e}_2(\theta))V - (1 + \lambda)\psi_1(\hat{e}_1(\theta), \theta_1) = R(\hat{e}_1(\theta))$$

$$\pi_2(\hat{e}_1(\theta), \hat{e}_2(\theta))V - (1 + \lambda)\psi_1(\hat{e}_2(\theta), \theta_1) = 0$$

where $R(\hat{e}_2) = 0$ simply because, given $\alpha = 1$, we have $m_2(\theta_1, \theta_2) = \theta_1$ and therefore $\partial m_2/\partial \theta_2 = 0$. When we also take into account the strategic substitutability of efforts ($\pi_{12} < 0$), we finally obtain a downward distortion in the equilibrium effort of the highest-type firm and an upward distortion in the equilibrium effort of the lowest-type firm. In other words, the regulator optimally substitutes the costly (in terms of informational rents) effort of the most skilled innovator for the costless (also in terms of informational rents) effort of the least skilled innovator. We call it the *rent-saving effect*. As the capacity of firms to assimilate the skills of their partner decreases (i.e., as α decreases), the regulator needs to increase the rents of the low-type (formally $R(\hat{e}_2) \propto \partial m_2/\partial \theta_2 = 1 - \alpha$). Therefore, the rent-saving effect decreases, and an upward effort distortion becomes less likely.

The size of the rent-saving effect depends crucially on the distribution of types. Suppose we increase the fraction of low-type firms in the economy (formally, θ is drawn from distribution $G(\cdot)$ instead of $F(\cdot)$ as stated in Definition 3). Since fewer firms have a high-type, then for a given θ_i the marginal cost of informational rents is smaller (formally, $\frac{1-G(\theta_i)}{g(\theta_i)} < \frac{1-F(\theta_i)}{f(\theta_i)} \Leftrightarrow R(e_i \,|\, G) < R(e_i \,|\, F)$). Thus, the rent-saving effect diminishes, so the regulator has incentives to distort less the efforts of both firms compared to the first-best.

The rent-saving effect implies a clear policy prescription. In cooperative projects where firms share homogeneous skills that cannot be observed, it is optimal to request more effort in the R&D stage from the innovator that

---

[25] Naturally, before revelation, firms do not know whether their type is greater or smaller than that of their partner so, in expectation, they always anticipate some informational rents.

turns out to be the least efficient. The idea is simply that inducing the most efficient firm to disclose its skills is costly, and this cost can be decreased only if little subsequent effort is required. Conversely, since the type of the (ex-post) least-efficient firms does not need to be learned, the regulator can costlessly require high efforts to compensate for its lower skills. This effect is relevant when firms come from the same industry and cooperate to develop a particular technology (as for SEMATECH and VLSI), and when projects aim at favoring the exchange of knowledge between ex-ante heterogeneous firms (e.g., the European Commission's policy of financing projects that involve firms from least favored regions).[26] Naturally, in order to induce the selection of these optimal efforts, the regulator must offer adequate disclosure-contingent transfers.

V(ii).   *Heterogeneous Skills*

We now turn to the case of multiple complementary skills that can be added and combined to increase the efficiency of the cooperative project. The sharing function of heterogeneous skills m_i is formalized with Assumption 7. Under truthful revelation of types, both innovators have the same final type $m_i = \beta\theta_1 + \theta_2$, where β captures the importance of firm 1's skills relative to firm 2's skills in the cooperation function. Also, since $m_1 = m_2$, the first-best effort of both firms is the same ($e_1^* = e_2^* = e^*$).[27]

Proposition 6. When skills are heterogeneous, there exists $\overline{\beta}(> 1)$ such that:

(i) if $\beta > \overline{\beta}$, then $\hat{e}_2 > e^*$ and $\hat{e}_1 < e^*$.
(ii) if $\beta < \overline{\beta}$, then $\hat{e}_2 < e^*$. Moreover, $\hat{e}_1 > e^*$ if and only if $\theta_1 \gg \theta_2$.

Proof: See Appendix 6.

When both skills are combined in the realization of the project, skill-sharing induces a positive externality on each innovator. If one type of skills is more valuable than the other (in our case, the skills of innovator 1 since $\beta > 1$), the regulator is forced to grant higher rents to that firm in order to induce truthful revelation of its information ($R(e_1) > R(e_2)$). To decrease the amount of rents, it is then optimal to distort downwards the effort of firm 1. As a compensation and given the strategic substitutability of efforts, the effort of firm 2 is distorted upwards. Naturally, the downward distortion of

---

[26] There are many opportunities for organizations from different countries to participate in European programmes. ESPRIT particularly favours the involvement of researchers from Central Europe and the Baltics, the Mediterranean region, and the newly independent states of the former Soviet Union (NIS).

[27] As a remark, the analysis could be extended to the case where $m_1 \neq m_2$. The effects we highlight in this section would still be present.

e_1 and the upward distortion of e_2 are proportional to the difference in importance between the skills of firm 1 and the skills of firm 2, which is captured with the parameter β. Last, part (ii) states that, if firms have a priori similar weights in the skill-sharing function, then efforts above the first-best level may occur only if the type of one firm is much greater than that of the other. The logic for this result relies on the strategic substitutability of effort and therefore it is very much in line with part (ii) of Proposition 5.

V(iii).   *Policy Implications*

The implications of our model are simple. First, and contrary to usual practices by policy-makers like the European Commission, transfers that promote cooperative ventures are effective only if they establish specific splitting rules of the payments among the team members, and take into account that informational rents need to be left. The current practice of the European Commission is to meet half of the costs of the industrial partners via the Framework Programme budget and to cover all the marginal costs (those incurred specifically from participating in a project) of research institutions. Funding also includes an overhead that participants are free to split. From a general perspective, our analysis suggests that the Commission should make payments contingent on both the skill of each participant and the socially efficient efforts. In particular, if the level of skills is not specifically rewarded, each firm finds it profitable to 'mimic the behavior of a firm with a lesser skill'. This translates into a smaller dissemination of knowledge among team members, and the regulator cannot benefit from the rent-saving effect to induce the desired levels of effort.[28]

Second, it is crucial that the regulator determine which team member has the most valuable skills (either because its input is more essential or because its knowledge is more developed) in order to encourage skill-sharing. We have shown that this can be achieved with an appropriate system of transfers (e.g., higher payoffs for members who prove their higher relative value). For homogeneous skills, the task is easy since skills belong to the same unique dimension, and the most valuable skill is simply the highest. In the case of SEMATECH and other projects that aim at helping firms catch up with technologies, the participants that turn out to have the most up-to-date knowledge after disclosure need to receive (ex-post) higher payments. In the case of European projects, it means in particular that the European Commission should not be concerned exclusively by equity issues when designing projects between the most and least favored regions. An optimal

---

[28] For instance, it is easy to see that even if firms share efficiently their skills, a transfer equal to a fraction of the total cost generates the same (suboptimal) effort from both firms in the case of homogeneous skills. If on top of that, they do not share their skills efficiently, even higher distortions are obtained.

transfer of technologies from research laboratories in the most favored regions (which are likely to be more efficient) to research laboratories in the least favored regions is possible if and only if the former are given explicit monetary incentives to disclose their skills. Paradoxically, if the goal is to put firms on equal footing, the means is to 'discriminate against' the least knowledgeable.

Naturally, identifying the most valuable skills is more complex when skills are heterogeneous. Indeed, skills now belong to different dimensions and essential skills are the skills in the most valuable dimension. Then, the regulator must identify this dimension *ex-ante*. Sometimes, fundamental research constitutes the crucial piece of the puzzle. This might be true for projects like ISS and HUMANE GENOME, and more generally when projects bring together leading scientists in fundamental research and engineers from the private sector. Skills necessary to develop innovations (e.g., develop new commercial products, medicines, etc.) might be less crucial because the task is similar to others already performed in related applications.[29] Sometimes, an area of expertise can benefit from the experience in another area to develop its own research (implications from biology on chemistry and vice-versa). The most valuable area of expertise will then be the area that is necessary not only to achieve the project but also to help researchers in other areas to make discoveries they would not make otherwise.[30] This effect is likely to be present in high-tech research programs that coordinate efforts in different fields.[31]

Third, the regulator must demand more effort from the least valuable members than from the most valuable ones to compensate for their lower skills. In the case of projects involving homogeneous skills, a low skill firm must dedicate more time, effort and personnel to accomplish the project. In the case of heterogeneous skills, participants who contribute less to build the overall initial capital of knowledge must also invest higher resources.

Overall, note that the results are qualitatively the same in the homogeneous and heterogeneous skills cases. When the skill of an innovator is relatively more important than the skill of the other (because it is better ranked as in the homogenous case, or because it is more essential as in the

---

[29] Although necessary, some routines are likely not to be crucial for the project (e.g., storing the information on genes and sequences in databases for the HUMAN GENOME program).

[30] Many scientific achievements follow this pattern. Understanding and combating zoonoses (transmissible animal diseases) has been possible by bringing research in veterinary sciences to human medicine. In the same lines, genetics benefits other areas of medicine as it is playing an increasingly important role in the diagnosis, monitoring, and treatment of diseases.

[31] For instance, the ISS monitors the project, 'Space Flight-Induced Reactivation of the Epstein-Barr Virus' aiming at studying a series of latent viruses. Scientists from NASA observed that such viruses can reactivate during space travel, possibly posing serious health problems. Coordinating the efforts of researchers in immunology and space travels, the experiment is expected to provide important information that may lead to a better understanding of latent herpes virus reactivation in humans living on Earth.

heterogenous case), the regulator substitutes the effort of the innovator with the most valuable skill with the effort of the innovator with the least valuable skill. The main difference between the two situations is that, in the homogeneous case, the identity of the most valuable innovator is endogenous and learned only ex-post: it depends on the privately known skill levels. By contrast, in the heterogeneous case, the identity of the most valuable innovator is common knowledge, and depends on exogenous factors such as the relative importance of the different areas of expertise for the success of the project.

VI. CONCLUDING REMARKS

The promotion of cooperative R&D programs has become a major tool of industrial policy both within countries and at the international level. From a general perspective, the difficulty for a regulator in organize research programs lies in her inability to observe the firms' technologies and skills and to monitor their decisions. Even though the inefficiencies due to incomplete information have been largely studied in the economics literature, few steps are taken in practice to correct them.

Conscious of the existing asymmetries of information, policy-makers offer schemes to reduce them. For instance, participants have to identify in detail the expected impacts of the results and the timescale on which these may be industrially implemented. However, the main criticism that is levelled against public intervention is the tendency to reward firms without monitoring their activities. For instance, the regulator offers equal amounts of money to each participant. Also, a major form of financing is the system of shared cost actions: when the European Commission selects a cooperative project after a call for proposals, she finances 50% of the costs without specifying strict rules on how these payments must be allocated between the team members. As has become clear in regulation theory, efficient skill-sharing and optimal efforts are provided if and only if the regulator selects adequate transfers.

As this paper shows, public authorities must select the payments that encourage the revelation of private information and the selection of socially optimal efforts. Since disclosure of skills is costly, the regulator distorts R&D efforts relative to first-best levels. When efforts are strategic complements, both efforts are distorted downwards. When efforts are strategic substitutes, the effort of the innovator with most valuable skills is distorted downwards (to reduce its rents) and that of its partner may be distorted upwards (to compensate for the efficiency loss). Overall, contrary to the standard non-cooperative regulation analyses, we predict that efforts are not necessarily lower under asymmetric information than under complete information.

Finally, we would like to point out one avenue for future research. In many situations, innovators are involved in R&D projects for which they are

partners at some stages and rivals at others. It could be of interest to analyze the incentives to disclose valuable information in a cooperative project when firms become rival in future R&D projects.[32] Intuitively, informational rents should be given not only to avoid a given firm's mimic in the behavior of a partner with a less valuable skill, but also to compensate its future utility loss generated by the use of that skill by second-generation rivals.

REFERENCES

Arrow, K., 1962, 'Economic Welfare and the Allocation of Resources for Invention,' in R.R. Nelson (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors* (Princeton University Press) pp. 609–626.

Arrow, K., 1979, 'The Property Rights Doctrine and Demand Revelation under Incomplete Information,' in *Economies and Human Welfare*, (Academic Press).

Baron, D. and Myerson, R. B., 1982, 'Regulating a Monopolist with Unknown Costs,' *Econometrica*, 50, pp. 911–930.

Bhattacharya, S.; Glazer, J. and Sappington, M., 1990, 'Sharing Productive Knowledge in Internally Financed R&D Contests,' *Journal of Industrial Economics*, 39, pp. 187–208.

Bhattacharya, S.; Glazer, J. and Sappington, M., 1992, 'Licensing and the Sharing of Knowledge in Research Joint Ventures,' *Journal of Economic Theory*, 56, pp. 43–69.

Brocas, I., 2002, 'Multi-stage Contracting with Applications to R&D and Insurance Policies,' mimeo, Columbia.

D'Aspremont, C.; Bhattacharya, S. and Gérard-Varet, L. -A., 1998, 'Knowledge as a Public Good: Efficient Sharing and Incentives for Development Effort,' *Journal of Mathematical Economics*, 30, pp. 389–404.

D'Aspremont, C. and Gérard-Varet, L. -A., 1979, 'Incentives and Incomplete Information,' *Journal of Public Economics*, 11, pp. 25–45.

D'Aspremont, C. and Jacquemin, A., 1988, 'Cooperative and Noncooperative R&D in Duopoly with Spillovers,' *American Economic Review*, 78, pp. 1133–1137.

Gandal, N. and Scotchmer, S., 1993, 'Coordinating Research Through Research Joint Ventures,' *Journal of Public Economics*, 51, pp. 173–193.

Hagedoorn, J., 1993, 'Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modes of Cooperation and Sectoral Differences,' *Strategic Management Journal*, 14, pp. 37–385.

Hamel, G., 1991, 'Competition for Competence and Inter-partner Learning within International Strategic Alliances,' *Strategic Management Journal*, 12, pp. 83–103.

Holmström, B., 1982, 'Moral Hazard in Teams,' *The Bell Journal of Economics*, pp. 325–340.

Jullien, B., 2000, 'Participation Constraints in Adverse Selection Models,' *Journal of Economic Theory*, 93, pp. 1–47.

Katz, M. L., 1986, 'An Analysis of Cooperative Research and Development,' *RAND Journal of Economics*, 17, pp. 527–543.

Katz, M. L. and Ordover, J. A., 1990, 'R&D Cooperation and Competition,' *Brookings Papers on Economic Activity: Microeconomics*, pp. 137–203.

---

[32] A related issue has been studied by Perez-Castrillo and Sandonis [1996]. The authors analyze RJV contracts under the assumption that the know-how is not contractible when partners are competitors on other markets. They show that partners have little incentive to share their know-how. As a result, profitable RJVs are sometimes not started.

Laffont, J. J. and Tirole, J., 1986, 'Using Cost Observation to Regulate Firms,' *Journal of Political Economy*, 94, pp. 614–641.

Laffont, J. J. and Tirole, J., 1990, 'The Regulation of Multiproduct Firms I: Theory,' *Journal of Public Economics*, 43, pp. 1–36.

Milgrom, P. and Roberts, J., 1990, 'The Economics of Modern Manufacturing: Technology, Strategy, and Organization,' *American Economic Review*, 80, pp. 511–528.

Perez-Castrillo, J. D. and Sandonis, J., 1997, 'Disclosure of Know-how in Research Joint Ventures,' *International Journal of Industrial Organization*, 15, pp. 51–75.

Sakakibara, M., 1997, 'Heterogeneity of Firm Capabilities and Cooperative Research and Development: An Empirical Examination of Motives,' *Strategic Management Journal*, 18, pp. 143–164.

APPENDIX

Proof of Proposition 1

Given W and $u_i(\theta) = 0$, innovator 1's optimal effort $e_1^{FB}(e_2, \theta_1, \theta_2, \lambda)$ is such that:

$$(1) \qquad \pi_1(e_1^{FB}(e_2, \theta_1, \theta_2, \lambda), e_2)V - (1+\lambda)\psi_1(e_1^{FB}(e_2, \theta_1, \theta_2, \lambda), m_1(\theta_1, \theta_2)) = 0.$$

We can differentiate (1) with respect to e_2, θ_1, θ_2 and λ:

$$[\pi_{11}(e_1, e_2)V - (1+\lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{FB}}{\partial e_2} + \pi_{12}(e_1, e_2)V = 0,$$

$$[\pi_{11}(e_1, e_2)V - (1+\lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{FB}}{\partial \theta_1} - (1+\lambda)\psi_{12}(e_1, m_1(\theta))\frac{\partial m_1}{\partial \theta_1} = 0,$$

$$[\pi_{11}(e_1, e_2)V - (1+\lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{FB}}{\partial \theta_2} - (1+\lambda)\psi_{12}(e_1, m_1(\theta))\frac{\partial m_1}{\partial \theta_2} = 0,$$

$$[\pi_{11}(e_1, e_2)V - (1+\lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{FB}}{\partial \lambda} - \psi_1(e_1, m_1(\theta)) = 0.$$

Given Assumption 8 and noting that the optimal effort of innovator 2 is a function $e_2^{FB}(e_1, \theta, \lambda)$ that satisfies the symmetric first-order condition to (1), we have:

- $\frac{\partial e_1^{FB}}{\partial e_2} \gtrless 0$ and $\frac{\partial e_2^{FB}}{\partial e_1} \gtrless 0$ if $\pi_{12} \gtrless 0$;

- $\frac{\partial e_1^{FB}}{\partial \theta_1} > 0, \frac{\partial e_1^{FB}}{\partial \theta_2} > 0, \frac{\partial e_2^{FB}}{\partial \theta_2} > 0$ and $\frac{\partial e_2^{FB}}{\partial \theta_1} > 0$. Given Assumptions 6 and 7, $\frac{\partial m_1}{\partial \theta_1} \gtrless \frac{\partial m_2}{\partial \theta_1}$ so $\frac{\partial e_1^{FB}}{\partial \theta_1} \gtrless \frac{\partial e_2^{FB}}{\partial \theta_1}$.

- $\frac{\partial e_1^{FB}}{\partial \lambda} < 0$ and $\frac{\partial e_2^{FB}}{\partial \lambda} < 0$.

A sufficient condition to have a unique equilibrium is $\left|\frac{\partial e_1^{FB}}{\partial e_2}\right| < 1$. Note that $\left|\frac{\partial e_1^{FB}}{\partial e_2}\right| = \left|\frac{-\pi_{12}(e_1, e_2)V}{\pi_{11}(e_1, e_2)V - (1+\lambda)\psi_{11}(e_1, m_1(\theta))}\right|$. In particular, the solution is unique if $\pi_{11}(e_1, e_2) < -|\pi_{12}(e_1, e_2)|$.

The equilibrium efforts are such that $e_1^*(\theta) = e_1^{FB}(e_2^*(\theta), \theta)$ and $e_2^*(\theta) = e_2^{FB}(e_1^*(\theta), \theta)$. Hence:

$$\frac{\partial e_1^*}{\partial \theta_1}\left[1 - \frac{\partial e_1^{FB}}{\partial e_2}\bigg|_{e_2^*}\frac{\partial e_2^{FB}}{\partial e_1}\bigg|_{e_1^*}\right] = \frac{de_1^{FB}}{de_2}\bigg|_{e_2^*}\frac{\partial e_2^{FB}}{\partial \theta_1} + \frac{\partial e_1^{FB}}{\partial \theta_1}$$

Given our previous results, we know that $1 - \frac{\partial e_1^{FB}}{\partial e_2}\big|_{e_2^*}\frac{\partial e_2^{FB}}{\partial e_1}\big|_{e_1^*} > 0$ and $\frac{de_1^{FB}}{de_2}\big|_{e_2^*}\frac{\partial e_2^{FB}}{\partial \theta_1}$ $+\frac{\partial e_1^{FB}}{\partial \theta_1} > 0$. Then $\frac{\partial e_1^*}{\partial \theta_1} > 0$ and $\frac{\partial e_2^*}{\partial \theta_2} > 0$. We also have:

$$\frac{\partial e_1^*}{\partial \lambda}\left[1 - \frac{\partial e_1^{FB}}{\partial e_2}\bigg|_{e_2^*}\frac{\partial e_2^{FB}}{\partial e_1}\bigg|_{e_1^*}\right] = \frac{de_1^{FB}}{de_2}\bigg|_{e_2^*}\frac{\partial e_2^{FB}}{\partial \lambda} + \frac{\partial e_1^{FB}}{\partial \lambda}$$

Then, $\frac{\partial e_1^*}{\partial \lambda} < 0$.

Proof of Lemma 1

In the third stage, for any vector of reports $\tilde{\theta}$, the optimal effort of firm i is:

$$\tilde{e}_i(\tilde{\theta}, \tau_i) = \arg\max_{e_i} \int_{\underline{\theta_j}}^{\overline{\theta_j}} [\pi(e_i, \tilde{e}_j(\tilde{\theta}, \tau_j))t_1^S(\tilde{\theta}) + (1 - \pi(e_i, \tilde{e}_j(\tilde{\theta}, \tau_j)))t_1^F(\tilde{\theta})$$
$$- \psi(e_i, m_i(\tau_i, \tilde{\theta}_j))]d\tilde{F}_j(\theta_j|\tilde{\theta}_j)$$

where $\tilde{F}_j(\theta_j|\tilde{\theta}_j)$ is such that the probability that $\theta_j = \tilde{\theta}_j(= \tau_j)$ is equal to 1 (it is a degenerate distribution). As a consequence, for all $\tilde{\theta}$:

$$\tilde{e}_i(\tilde{\theta}, \tau_i) = \arg\max_{e_i} \pi(e_i, \tilde{e}_j(\tilde{\theta}, \tilde{\theta}_j))t_1^S(\tilde{\theta}) + (1 - \pi(e_i, \tilde{e}_j(\tilde{\theta}, \tilde{\theta}_j)))t_1^F(\tilde{\theta})$$
$$- \psi(e_i, m_i(\tau_i, \tilde{\theta}_j))$$

as stated in (MH). In a perfect Bayesian Nash equilibrium, firm i anticipates that its partner selects $\sigma(\theta_j) = \theta_j$. Then, the vector of reported types $\tilde{\theta}$ is simply $(\tilde{\theta}_i, \theta_j)$. Let us determine the conditions under which $\sigma(\theta_i) = \theta_i$. First, given that $\tau_i = \max(\theta_i, \tilde{\theta}_i)$, we can rewrite the expected utility of firm i as:

$$\Phi_i(\tilde{\theta}_i, \theta_i) = \begin{cases} \Phi_i^-(\tilde{\theta}_i, \theta_i) & \text{if} \quad \tilde{\theta}_i \leqslant \theta_i \\ \Phi_i^+(\tilde{\theta}_i, \theta_i) & \text{if} \quad \tilde{\theta}_i \geqslant \theta_i \end{cases}$$

where:

$$\Phi_i^-(\tilde{\theta}_i, \theta_i) = E_{\theta_j}[t_i^S(\tilde{\theta}_i, \theta_j)\pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \theta_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))$$
$$+ t_i^F(\tilde{\theta}_i, \theta_j)[1 - \pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \theta_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))]$$
$$- \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))]$$

$$\Phi_i^+(\tilde{\theta}_i, \theta_i) = E_{\theta_j}[t_i^S(\tilde{\theta}_i, \theta_j)\pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))$$
$$+ t_i^F(\tilde{\theta}_i, \theta_j)[1 - \pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))]$$
$$- \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\tilde{\theta}_i, \tilde{\theta}_i))] - c(\theta_i, \tilde{\theta}_i)$$

Consider two types $\tilde{\theta}_i$ and θ_i such that $\tilde{\theta}_i < \theta_i$. An agent with type θ_i reports truthfully if $\Phi_i(\theta_i, \theta_i) \geqslant \Phi_i(\tilde{\theta}_i, \theta_i)$ or, equivalently, if $\Phi_i(\theta_i, \theta_i) \geqslant \Phi_i(\tilde{\theta}_i, \tilde{\theta}_i) +$

$[\Phi_i^-(\tilde{\theta}_i, \theta_i) - \Phi_i^-(\tilde{\theta}_i, \tilde{\theta}_i)]$. An agent with type $\tilde{\theta}_i$ reports truthfully if $\Phi_i(\tilde{\theta}_i, \theta_i) \geqslant \Phi_i(\theta_i, \theta_i)$ or, $\Phi_i(\tilde{\theta}_i, \theta_i) \geqslant \Phi_i(\theta_i, \theta_i) + [\Phi_i^+(\theta_i, \theta_i) - \Phi_i^+(\theta_i, \theta_i)]$. Overall, we must have:

$$\Phi_i^-(\tilde{\theta}_i, \theta_i) - \Phi_i^-(\tilde{\theta}_i, \tilde{\theta}_i) \leqslant \Phi_i(\theta_i, \theta_i) - \Phi_i(\tilde{\theta}_i, \tilde{\theta}_i) \leqslant \Phi_i^+(\theta_i, \theta_i) - \Phi_i^+(\theta_i, \tilde{\theta}_i)$$

and a necessary condition for truthtelling is simply:

$$(2) \qquad \Phi_i^-(\tilde{\theta}_i, \theta_i) - \Phi_i^-(\tilde{\theta}_i, \tilde{\theta}_i) \leqslant \Phi_i^+(\theta_i, \theta_i) - \Phi_i^+(\theta_i, \tilde{\theta}_i)$$

The innovator chooses its effort so as to maximize its expected utility. In particular, if i's type is θ_i and its report $\tilde{\theta}_i$, i prefers to exert $\tilde{e}_i(\tilde{\theta}_i, \theta_j, \theta_i)$ rather than $\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i)$:

$$\pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \theta_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))[t_i^S(\tilde{\theta}_i, \theta_j) - t_i^F(\tilde{\theta}_i, \theta_j)] + t_i^F(\tilde{\theta}_i, \theta_j)$$

$$- \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j)) \geqslant \pi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), \tilde{e}_j(\tilde{\theta}_i, \theta_j, \theta_j))[t_i^S(\tilde{\theta}_i, \theta_j)$$

$$- t_i^F(\tilde{\theta}_i, \theta_j)] + t_i^F(\tilde{\theta}_i, \theta_j) - \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\theta_i, \theta_j))$$

Using this inequality, we get that:

$$\Phi_i^-(\tilde{\theta}_i, \theta_i) - \Phi_i^-(\tilde{\theta}_i, \tilde{\theta}_i) \geqslant E_{\theta_j}\Big[\psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\theta_i, \theta_j))\Big]$$

Moreover, $\Phi_i^+(\theta_i, \theta_i) - \Phi_i^+(\theta_i, \tilde{\theta}_i) = c(\tilde{\theta}_i, \theta_i)$. This can be rewritten as:

$$\Phi_i^+(\theta_i, \theta_i) - \Phi_i^+(\theta_i, \tilde{\theta}_i) = c(\tilde{\theta}_i, \theta_i) + E_{\theta_j}[\psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))$$
$$- \psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))]$$

Since $m_i(\theta_i, \theta_j) > m_i(\tilde{\theta}_i, \theta_j)$ and $\psi_2(\cdot, \cdot) < 0$, we have:

$$\Phi_i^+(\theta_i, \theta_i) - \Phi_i^+(\theta_i, \tilde{\theta}_i) \leqslant c(\tilde{\theta}_i, \theta_i) + E_{\theta_j}[\psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\tilde{\theta}_i, \theta_j))$$
$$- \psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))]$$

Then, (2) becomes:

$$E_{\theta_j}[\psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\theta_i, \theta_j))]$$

$$\leqslant c(\tilde{\theta}_i, \theta_i) + E_{\theta_j}[\psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))]$$

It is true for all $\tilde{\theta}_i < \theta_i$ if $\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i) \leqslant \tilde{e}_i(\theta_i, \theta_j, \theta_i)$ for all $\tilde{\theta}_i < \theta_i$. Overall, a necessary condition for truthtelling is (IC₂). Besides, for all θ_i and $\tilde{\theta}_i = \theta_i - \delta$:

$$E_{\theta_j}\Big[\psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\theta_i, \theta_j))\Big]$$

$$\leqslant \Phi(\theta_i, \theta_i) - \Phi(\tilde{\theta}_i, \tilde{\theta}_i)$$

$$\leqslant c(\tilde{\theta}_i, \theta_i) + E_{\theta_j}\Big[\psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))\Big]$$

and if $\delta \to 0$, we must have

$$E_{\theta_j}\left[\psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\tilde{\theta}_i, \theta_j, \tilde{\theta}_i), m_i(\theta_i, \theta_j))\right]$$

$$\leqslant \Phi(\theta_i, \theta_i) - \Phi(\tilde{\theta}_i, \tilde{\theta}_i)$$

$$\leqslant E_{\theta_j}\left[\psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\tilde{\theta}_i, \theta_j)) - \psi(\tilde{e}_i(\theta_i, \theta_j, \theta_i), m_i(\theta_i, \theta_j))\right]$$

Since $\tilde{e}_i(\theta_i, \theta_j, \theta_i) \equiv \tilde{e}_i(\theta_i, \theta_j)$ is increasing, we can take the Riemann integral. Then, the agent reveals truthfully if:

$$U_i(\theta_i) - U_i(\tilde{\theta}_i) = -\int_{\tilde{\theta}_i}^{\theta_i} E_{\theta_j}\left[\psi_2(\tilde{e}_i(s, \theta_j), m_i(s, \theta_j)\frac{\partial m_i}{\partial s}\right]ds$$

that is,

$$\frac{dU_i}{d\theta_i}(\theta_i) = -E_{\theta_j}\left[\psi_2(\tilde{e}_i(\theta), m_i(\theta))\frac{\partial m_i}{\partial \theta_i}\right]$$

which is (IC$_1$) in Lemma 1. So far, we have shown that $\Phi_i(\theta_i, \theta_i) \leqslant \Phi_i(\tilde{\theta}_i, \theta_i)$ for all θ_i and $\tilde{\theta}_i$ implies (IC$_1$) and (IC$_2$). We need to show that (IC$_1$) and (IC$_2$) also imply $\Phi_i(\theta_i, \theta_i) \geqslant \Phi_i(\tilde{\theta}_i, \theta_i)$ for all θ_i and $\tilde{\theta}_i$. To show this, consider $\tilde{\theta}_i < \theta_i$ and assume $\Phi_i^-(\tilde{\theta}_i, \theta_i) > \Phi_i^-(\theta_i, \theta_i)$. Then we have, $\int_{\theta_i}^{\tilde{\theta}_i}\frac{\partial}{\partial s}\Phi_i^-(s, \theta_i)ds > 0$. Using (IC$_1$) and the fact that firm i chooses its effort so as to maximize its utility, we get that $\frac{\partial}{\partial \theta_i}\Phi_i(\tilde{\theta}_i, \theta_i)|_{\tilde{\theta}_i = \theta_i} = 0$. Then, the last inequality can be rewritten as $\int_{\theta_i}^{\tilde{\theta}_i}[\frac{\partial}{\partial s}\Phi_i^-(s, \theta_i) - \frac{\partial}{\partial s}\Phi_i^-(s, s)ds] > 0$, or $\int_{\theta_i}^{\tilde{\theta}_i}\frac{\partial}{\partial s}\frac{\partial}{\partial t}\Phi_i^-(s, t)dt\, ds > 0$. Given that i chooses its effort so as to maximize its expected utility, we have

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t}\Phi_i^-(s, t) = -E_{\theta_j}\frac{\partial}{\partial s}\tilde{e}_i(s, t)\psi_{12}(\tilde{e}_i(s, t), m_i(s, t))\frac{\partial}{\partial s}m_i(s, t)dt.$$

(IC$_2$) implies that this term is positive, which leads to a contradiction for all $\tilde{\theta}_i < \theta_i$.

Consider now $\tilde{\theta}_i > \theta_i$. We have $\Phi_i(\tilde{\theta}_i, \theta_i) = \Phi(\tilde{\theta}_i, \tilde{\theta}_i) + c(\tilde{\theta}_i, \tilde{\theta}_i) - c(\theta_i, \tilde{\theta}_i)$. Using (IC$_1$), we have $\Phi_i(\tilde{\theta}_i, \theta_i) = U(\theta_i) - \int_{\theta_i}^{\tilde{\theta}_i} E_{\theta_j}[\psi_2(\tilde{e}_i(s, \theta_j), m_i(s, \theta_j))\frac{\partial m_i}{\partial s}]ds + c(\tilde{\theta}_i, \tilde{\theta}_i) - c(\theta_i, \tilde{\theta}_i)$. Using (IC$_2$), we get

$$\Phi_i(\tilde{\theta}_i, \theta_i)U(\theta_i) - (\tilde{\theta}_i - \theta_i)\psi_2(\tilde{e}_i(\tilde{\theta}_i, \theta_j), m_i(\tilde{\theta}_i, \theta_j))\frac{\partial m_i}{\partial \tilde{\theta}_j} + c(\tilde{\theta}_i, \tilde{\theta}_i) - c(\theta_i, \tilde{\theta}_i)$$

$$= U(\theta_i) + (\tilde{\theta}_i - \theta_i)\left[-\psi_2(\tilde{e}_i(\tilde{\theta}_i, \theta_j), m_i(\tilde{\theta}_i, \theta_j))\frac{\partial m_i}{\partial \tilde{\theta}_i} + \frac{c(\tilde{\theta}_i, \tilde{\theta}_i) - c(\theta_i, \tilde{\theta}_i)}{\tilde{\theta}_i - \theta_i}\right]$$

Note that for all $\tilde{\theta}_i > \theta_i$, there exists θ^* such that $c(\theta_i, \tilde{\theta}_i) = -c_1(\theta^*, \tilde{\theta}_i)(\tilde{\theta}_i - \theta_i) + c(\tilde{\theta}_i, \tilde{\theta}_i)$. Given assumption 9, we have $\Phi_i(\tilde{\theta}_i, \theta_i) \leqslant U(\theta_i)$.

To complete the proof, note that if (IR$_{ep}$) is satisfied, then (IR$_{ea}$) is also satisfied. Last, the expression of the ex post welfare is the same as under complete information. Given incomplete information, the seller maximizes the expected welfare. $\square$

Proof of Proposition 2

Given W', if the types of firms are (θ_1, θ_2), then the regulator wants to induce innovator 1 to select the effort $e_1^{SB}(e_2, \theta_1, \theta_2)$ that satisfies:

$$
(3) \quad
\begin{aligned}
&\pi_1(e_1^{SB}(e_2, \theta), e_2)V - (1 + \lambda)\psi_1(e_1^{SB}(e_2, \theta), m_1(\theta)) \\
&+ \lambda\psi_{12}(e_1^{SB}(e_2, \theta), m_1(\theta))\frac{1 - F_1(\theta_1)}{f_1(\theta_1)}\frac{dm_1}{d\theta_1} = 0
\end{aligned}
$$

Differentiating (3) with respect to e_2, θ_1 and θ_2, we get:

$$
[\pi_{11}(e_1, e_2)V - (1 + \lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{SB}}{\partial e_2} + \pi_{12}(e_1, e_2)V = 0,
$$

$$
\begin{aligned}
&[\pi_{11}(e_1, e_2)V - (1 + \lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{SB}}{\partial \theta_1} \\
&- \psi_{12}(e_1, m_1(\theta))\frac{\partial m_1}{\partial \theta_1}\left[1 + \lambda - \lambda\frac{d}{d\theta_1}\left(\frac{1 - F_1(\theta_1)}{f_1(\theta_1)}\right)\right] = 0,
\end{aligned}
$$

$$
\begin{aligned}
&[\pi_{11}(e_1, e_2)V - (1 + \lambda)\psi_{11}(e_1, m_1(\theta))]\frac{\partial e_1^{SB}}{\partial \theta_2} \\
&- (1 + \lambda)\psi_{12}(e_1, m_1(\theta))\frac{\partial m_1}{\partial \theta_2} = 0.
\end{aligned}
$$

From (1) and (3) and given $\psi_{12} < 0$, it is immediate that $e_1^{SB}(e_2, \overline{\theta}_1, \theta_2) = e_1^{FB}(e_2, \overline{\theta}_1, \theta_2)$ and $e_1^{SB}(e_2, \theta_1, \theta_2) < e_1^{FB}(e_2, \theta_1, \theta_2)$ for all $\theta_1 < \overline{\theta}_1$. By symmetry, $e_2^{SB}(e_1, \theta_1, \overline{\theta}_2) = e_2^{FB}(e_1, \theta_1, \overline{\theta}_2)$ and $e_2^{SB}(e_1, \theta_1, \theta_2) < e_2^{FB}(e_1, \theta_1, \theta_2)$. Also, using the same reasoning as in Proposition 1:

- $\frac{\partial e_1^{SB}}{\partial e_2} \gtrless 0$ and $\frac{\partial e_2^{SB}}{\partial e_1} \gtrless 0$ if $\pi_{12} \gtrless 0$;

- $\frac{\partial e_1^{SB}}{\partial \theta_1} > 0$, $\frac{\partial e_1^{SB}}{\partial \theta_2} > 0$, $\frac{\partial e_2^{SB}}{\partial \theta_2} > 0$, $\frac{\partial e_2^{SB}}{\partial \theta_1} > 0$ and $\frac{\partial e_1^{SB}}{\partial \theta_1} \geqslant \frac{\partial e_2^{SB}}{\partial \theta_1}$.

The equilibrium efforts are such that $\hat{e}_1(\theta) = e_1^{SB}(\hat{e}_2(\theta), \theta)$ and $\hat{e}_2(\theta) = e_2^{SB}(\hat{e}_1(\theta), \theta)$. Using the same reasoning as in Appendix 1, it is immediate that $\frac{\partial \hat{e}_1}{\partial \theta_1} > 0$ and $\frac{\partial \hat{e}_2}{\partial \theta_2} > 0$. This means that the efforts $\{\hat{e}_1, \hat{e}_2\}$ that maximize W' also satisfy the constraint (IC$_2$). Since this is the only remaining constraint of the optimization program, we have proved that $\hat{e}_1$ and $\hat{e}_2$ are the optimal solutions to P.

The only issue left is to determine the transfers $\hat{t}_i^S(\theta)$ and $\hat{t}_i^F(\theta)$ that implement the optimal contract. If both innovators reveal truthfully, they select the efforts $(\tilde{e}_1(\theta), \tilde{e}_2(\theta))$ that maximize their utility, as given by (MH). For a given pair $\{\hat{t}_i^S(\theta), \hat{t}_i^F(\theta)\}$, innovator i's effort $\tilde{e}_i(\theta)$ is:

$$
\pi_i(\tilde{e}_i(\theta), \tilde{e}_j(\theta))[\hat{t}_i^S(\theta) - \hat{t}_i^F(\theta)] - \psi_1(\tilde{e}_i(\theta), m_i(\theta)) = 0
$$

This effort coincides with the desired level $\hat{e}_i(\theta)$ if and only if $\hat{t}_i^S(\theta) - \hat{t}_i^F(\theta) = \frac{\psi_1(\hat{e}_i(\theta), m_i(\theta))}{\pi_i(\hat{e}_i(\theta), \tilde{e}_j(\theta))}$. Transfers must also be such that the final utility of innovator i

coincides with the utility $\hat{u}_i(\theta)$ that the regulator wants to implement in the optimal second-best contract, which is given by:

$$\hat{u}_i(\theta) = -\int_{\underline{\theta}_i}^{\theta_i} \psi_2(\hat{e}_i(s, \theta_j), m_i(s, \theta_j)) \frac{\partial m_i}{\partial s} ds$$

Overall, combining

$$\hat{t}_i^S(\theta) - \hat{t}_i^F(\theta) = \frac{\psi_1(\hat{e}_i(\theta), m_i(\theta))}{\pi_i(\hat{e}_i(\theta), \hat{e}_j(\theta))}$$

and

$$\pi(\hat{e}_i(\theta), \hat{e}_j(\theta)) \hat{t}_i^S(\theta) + (1 - \pi(\hat{e}_i(\theta), \hat{e}_j(\theta))) \hat{t}_i^F(\theta) - \psi(\hat{e}_i, m_i(\theta)) = \hat{u}_i(\theta)$$

we conclude that the following pair of transfers:

$$\hat{t}_i^S(\theta) = \psi(\hat{e}_i(\theta), m_i(\theta)) + \frac{\psi_1(\hat{e}_i(\theta), m_i(\theta))}{\pi_i(\hat{e}_i(\theta), \hat{e}_j(\theta))} (1 - \pi(\hat{e}_i(\theta), \hat{e}_j(\theta))) + \hat{u}_i(\theta),$$

$$\hat{t}_i^F(\theta) = \psi(\hat{e}_i(\theta), m_i(\theta)) - \frac{\psi_1(\hat{e}_i(\theta), m_i(\theta))}{\pi_i(\hat{e}_i(\theta), \hat{e}_j(\theta))} \pi(\hat{e}_i(\theta), \hat{e}_j(\theta)) + \hat{u}_i(\theta) < \hat{t}_i^S(\theta)$$

implement the optimal mechanism. □

Proof of Propositions 3 and 4

We know that $e_i^{SB}(e_j, \theta_1, \theta_2) < e_i^{SB}(e_j, \theta_1, \theta_2)$ for all $\theta_1 < \overline{\theta}_1$ and $\theta_2 < \overline{\theta}_2$. The proof is then immediate once we note that:

- If $\pi_{12} > 0$, then $\frac{\partial e_i^{FB}}{\partial e_j} \in (0, 1)$ and $\frac{\partial e_i^{SB}}{\partial e_j} \in (0, 1)$.

- If $\pi_{12} < 0$, then $\frac{\partial e_i^{FB}}{\partial e_j} \in (0, 1)$ and $\frac{\partial e_i^{SB}}{\partial e_j} \in (-1, 0)$.

The two cases are illustrated in Figures 3 and 4. □

Proof of Proposition 5

(i) Let $\alpha = 1$. In that case and given $\pi(e^a, e^b) = \pi(e^b, e^a)$, the first-best reaction functions are $e_1^{FB}(e, \theta) = e_2^{FB}(e, \theta)$, so the first-best equilibrium efforts are symmetric $e_1^*(\theta) = e_2^*(\theta)$. Under incomplete information and if $\theta_1 > \theta_2$, we have $m_1(\theta) = m_2(\theta) = \theta_1$. Obviously, $\partial m_1/\partial \theta_1 > 0$ and $\partial m_2/\partial \theta_2 = 0$ and therefore $e_2^{SB}(e_1, \theta) = e_2^{FB}(e_1, \theta)$ and $e_1^{SB}(e_2, \theta) < e_1^{FB}(e_2, \theta)$. This immediately leads to $\hat{e}_2(\theta) > e_2^*(\theta)$ and $\hat{e}_1(\theta) < e_1^*(\theta)$.

(ii) Suppose now that $\alpha < 1$ and $\theta_1 > \theta_2$. To simplify notations, let $m_2 = \alpha\theta_1 + (1 - \alpha)\theta_2$.

- The first-best effort reaction functions of 1 and 2 are given by:

$$\pi_1(e_1^{FB}, e_2)V - (1 + \lambda)\psi_1(e_1^{FB}, \theta_1) = 0 \quad \text{and}$$

$$\pi_2(e_1, e_2^{FB})V - (1 + \lambda)\psi_1(e_2^{FB}, m_2) = 0$$

Differentiating the first-order conditions with respect to α, we get $\frac{\partial e_1^{FB}}{\partial \alpha} = 0$ and

$$[\pi_{22}(e_1, e_2^{FB})V - (1+\lambda)\psi_{11}(e_2^{FB}, m_2)]\frac{\partial e_2^{FB}}{\partial \alpha} - \psi_{12}(e_2^{FB}, m_2)(1+\lambda)(\theta_1 - \theta_2)$$

$$= 0 \Rightarrow \frac{\partial e_2^{FB}}{\partial \alpha} > 0$$

Therefore:

$$\frac{\partial e_1^*}{\partial \alpha}\left[1 - \frac{\partial e_1^{FB}}{\partial e_2}\bigg|_{e_2^*}\frac{\partial e_2^{FB}}{\partial e_1}\bigg|_{e_1^*}\right] = \frac{de_1^{FB}}{de_2}\bigg|_{e_2^*}\frac{\partial e_2^{FB}}{\partial \alpha} + \frac{\partial e_1^{FB}}{\partial \alpha} \Rightarrow \frac{\partial e_1^*}{\partial \alpha} < 0$$

$$\frac{\partial e_2^*}{\partial \alpha}\left[1 - \frac{\partial e_2^{FB}}{\partial e_1}\bigg|_{e_1^*}\frac{\partial e_1^{FB}}{\partial e_2}\bigg|_{e_2^*}\right] = \frac{de_2^{FB}}{de_1}\bigg|_{e_1^*}\frac{\partial e_1^{FB}}{\partial \alpha} + \frac{\partial e_2^{FB}}{\partial \alpha} \Rightarrow \frac{\partial e_2^*}{\partial \alpha} > 0$$

- Under asymmetric information, the first-order conditions are

$$\pi_1(e_1^{SB}, e_2)V - (1+\lambda)\psi_1(e_1^{SB}, \theta_1) + \lambda\psi_{12}(e_1^{SB}, \theta_1)\frac{1 - F(\theta_1)}{f(\theta_1)} = 0$$

$$\pi_2(e_1, e_2^{SB})V - (1+\lambda)\psi_1(e_2^{SB}, m_2) + \lambda\psi_{12}(e_2^{SB}, m_2)\frac{1 - F(\theta_2)}{f(\theta_2)}(1 - \alpha) = 0$$

Differentiating again the first-order conditions with respect to α, we get $\frac{\partial e_1^{SB}}{\partial \alpha} = 0$. Also, given $\psi_{122} = 0$ (see Assumption 8):

$$[\pi_{22}(e_1, e_2^{SB})V - (1+\lambda)\psi_{11}(e_2^{SB}, m_2)]\frac{\partial e_2^{SB}}{\partial \alpha} - \psi_{12}(e_2^{SB}, m_2)[(1+\lambda)(\theta_1$$

$$- \theta_2) + \lambda\frac{1 - F(\theta_2)}{f(\theta_2)}] = 0$$

and therefore $\frac{\partial e_2^{SB}}{\partial \alpha} > \frac{\partial e_2^{FB}}{\partial \alpha}$. Last, since

$$\frac{\partial \hat{e}_1}{\partial \alpha}\left[1 - \frac{\partial e_1^{SB}}{\partial e_2}\bigg|_{\hat{e}_2}\frac{\partial e_2^{SB}}{\partial e_1}\bigg|_{\hat{e}_1}\right] = \frac{de_1^{SB}}{de_2}\bigg|_{\hat{e}_2}\frac{\partial e_2^{SB}}{\partial \alpha} + \frac{\partial e_1^{SB}}{\partial \alpha}$$

$$\frac{\partial \hat{e}_2}{\partial \alpha}\left[1 - \frac{\partial e_2^{SB}}{\partial e_1}\bigg|_{\hat{e}_1}\frac{\partial e_1^{SB}}{\partial e_2}\bigg|_{\hat{e}_2}\right] = \frac{de_2^{SB}}{de_1}\bigg|_{\hat{e}_1}\frac{\partial e_1^{SB}}{\partial \alpha} + \frac{\partial e_2^{SB}}{\partial \alpha}$$

we immediately obtain that $\frac{\partial \hat{e}_1}{\partial \alpha} < \frac{\partial e_1^*}{\partial \alpha} < 0$ and $\frac{\partial \hat{e}_2}{\partial \alpha} > \frac{\partial e_2^*}{\partial \alpha} > 0$.

Overall there exists $\underline{\alpha}$ such that $\hat{e}_1(\underline{\alpha}) = e_1^*(\underline{\alpha})$ and $\hat{e}_2(\underline{\alpha}) < e_2^*(\underline{\alpha})$ and $\overline{\alpha}$ such that $\hat{e}_2(\overline{\alpha}) = e_2^*(\overline{\alpha})$ and $\hat{e}_1(\overline{\alpha}) < e_1^*(\overline{\alpha})$. If $\alpha < \underline{\alpha}$, then $\hat{e}_1(\alpha) > e_1^*(\alpha)$ and $\hat{e}_2(\alpha) < e_2^*(\alpha)$; if $\alpha \in (\underline{\alpha}, \overline{\alpha})$, then $\hat{e}_1(\alpha) < e_1^*(\alpha)$ and $\hat{e}_2(\alpha) < e_2^*(\alpha)$; if $\alpha > \overline{\alpha}$, then $\hat{e}_2(\alpha) > e_2^*(\alpha)$ and $\hat{e}_1(\alpha) < e_1^*(\alpha)$.

(iii) $\frac{d}{d\theta}\left(\frac{1 - F(\theta)}{1 - G(\theta)}\right) \propto -f(\theta)[1 - M(F(\theta)) - M'(F(\theta))(1 - F(\theta))] \propto -\frac{f(\theta)}{1 - F(\theta)} + \frac{g(\theta)}{1 - G(\theta)}$.

Moreover, $\frac{d}{d\theta}(1 - M(F(\theta)) - M'(F(\theta))(1 - F(\theta))) = -(1 - F(\theta))M''(F(\theta))f(\theta) > 0$ and $1 - M(F(\overline{\theta})) - M'(F(\overline{\theta}))(1 - F(\overline{\theta})) = 0$. Hence, $\frac{g(\theta)}{1 - G(\theta)} > \frac{f(\theta)}{1 - F(\theta)}$ and $\frac{1 - G(\theta)}{g(\theta)} < \frac{1 - F(\theta)}{f(\theta)}$ for all θ. This means that, when $\alpha = 1$ and $\theta_1 > \theta_2$ we have $e_1^{FB}(e_2, \theta) >$

$e_1^{SB}(e_2, \theta \mid G) > e_1^{SB}(e_2, \theta \mid F)$ and $e_2^{FB}(e_1, \theta) = e_2^{SB}(e_1, \theta \mid G) = e_2^{SB}(e_1, \theta \mid F)$ which leads to $\hat{e}_2^F > \hat{e}_2^G > e_2^*$ and $\hat{e}_1^F < \hat{e}_1^G < e_1^*$.

Proof of Proposition 6

1- Suppose first that $\beta = 1$. Under complete information the optimal effort of innovator i is a function $e_i^{FB}(e, \theta)$ of the effort of its partner such that

$$(4) \qquad \pi_i(e_i^{FB}(e_j, \theta), e_j)V - (1 + \lambda)\psi_1(e_i^{FB}(e_j, \theta), \theta_1 + \theta_2) = 0.$$

We have $e_1^{FB}(e, \theta) = e_2^{FB}(e, \theta)$ and in equilibrium, the first best efforts are symmetric. We denote the equilibrium effort by $e^*(\theta)$. Under asymmetric information, the optimal effort of innovator i is a function $e_i^{SB}(e_j, \theta)$ of the effort of its partner. The first order conditions are:

$$\pi_1(e_1^{SB}(e_2, \theta), e_2)V - (1 + \lambda)\psi_1(e_1^{SB}(e_2, \theta), \theta_1 + \theta_2)$$
$$+ \lambda\psi_{12}(e_1^{SB}(e_2, \theta), \theta_1 + \theta_2)\frac{1 - F(\theta_1)}{f(\theta_1)} = 0$$

$$\pi_2(e_1, e_2^{SB}(e_1, \theta))V - (1 + \lambda)\psi_1(e_2^{SB}(e_1, \theta), \theta_1 + \theta_2)$$
$$+ \lambda\psi_{12}(e_2^{SB}(e_1, \theta), \theta_1 + \theta_2)\frac{1 - F(\theta_2)}{f(\theta_2)} = 0$$

- When $\theta_1 = \theta_2$, the equilibrium is symmetric and the second best effort is $\hat{e}(\theta) \geqslant e^*(\theta)$. Given assumption 1, $e_1^{SB}(e, \theta) < e_2^{SB}(e, \theta)$ when $\theta_1 > \theta_2$ and $e_1^{SB}(e, \theta) > e_2^{SB}(e, \theta)$ when $\theta_1 < \theta_2$.
  We know from Appendix 1 and 3 that $e_1^{FB}(e, \theta)$, $e_2^{FB}(e, \theta)$, $e_1^{SB}(e, \theta)$ and $e_2^{SB}(e, \theta)$ are increasing in θ_1 and we have shown that both $e^*(\theta)$ and $\hat{e}_1(\theta)$ increase in θ_1.

- Under assumption 8, we have that $\frac{\partial e^{FB}}{\partial e_2} = \frac{\partial e^{SB}}{\partial e_2}$, $\frac{\partial e_2^{FB}}{\partial \theta_1} = \frac{\partial e_2^{SB}}{\partial \theta_1}$ and $\frac{\partial e_1^{FB}}{\partial \theta_1} < \frac{\partial e_1^{SB}}{\partial \theta_1}$. Then,
  $\frac{\partial \hat{e}_1}{\partial \theta_1} > \frac{\partial e^*}{\partial \theta_1}$. Similarly, $e^*(\theta)$ and $\hat{e}_2(\theta)$ increase in θ_2 and $\frac{\partial \hat{e}_2}{\partial \theta_2} > \frac{\partial e^*}{\partial \theta_2}$

As a consequence, there exists $\theta^b(\theta_2)$ such that $\hat{e}_1(\theta^b(\theta_2), \theta_2) = e^*(\theta^b(\theta_2), \theta_2)$. For all $\theta_1 > \theta^b(\theta_2)$, $\hat{e}_1(\theta) > e^*(\theta)$ and for all $\theta_1 < \theta^b(\theta_2)$, $\hat{e}_1(\theta) < e^*(\theta)$ Given that $\hat{e}_1(\theta_2, \theta_2) = \hat{e}_2(\theta_2, \theta_2) < e^*(\theta_2, \theta_2)$ and the fact that $\hat{e}_1(\theta)$ and $e^*(\theta)$ are increasing in θ_1, we have necessarily $\theta^b(\theta_2) > \theta_2$. Moreover,

$$\frac{\partial \theta^b}{\partial \theta_2}\left(\frac{\partial \hat{e}_1}{\partial \theta_1} - \frac{\partial e^*}{\partial \theta_1}\right) = \frac{\partial e^*}{\partial \theta_2} - \frac{\partial \hat{e}_1}{\partial \theta_2}$$

We have already shown that $\frac{\partial \hat{e}_1}{\partial \theta_1} > \frac{\partial e^*}{\partial \theta_1}$. Moreover,

$$\frac{\partial \hat{e}_1}{\partial \theta_2} = \frac{\partial e_1^{SB}}{\partial e_2}\frac{\partial \hat{e}_2}{\partial \theta_2} + \frac{\partial e_1^{SB}}{\partial \theta_2} = \frac{\partial e_1^{FB}}{\partial e_2}\frac{\partial \hat{e}_2}{\partial \theta_2} + \frac{\partial e_1^{FB}}{\partial \theta_2} < \frac{\partial e_1^{FB}}{\partial e_2}\frac{\partial e^*}{\partial \theta_2} + \frac{\partial e_1^{FB}}{\partial \theta_2} = \frac{\partial e^*}{\partial \theta_2}$$

Then $\frac{\partial \theta^b}{\partial \theta_2} > 0$. By symmetry, there also exists $\theta^a(\theta_1) > \theta_1$ such that $\hat{e}_2(\theta^a(\theta_1), \theta_1) = e^*(\theta^a(\theta_1), \theta_1)$ with $\frac{\partial \theta^a}{\partial \theta_1} > 0$. For all $\theta_2 > \theta^a(\theta_1)$, $\hat{e}_2(\theta) > e^*(\theta)$ and for all $\theta_2 < \theta^a(\theta_1)$, $\hat{e}_2(\theta) < e^*(\theta)$. To sum up, consider the functions $\theta^a(\theta_1)$ and

$\theta^{b^{-1}}(\theta_1) < \theta_1$. For all θ_1, (i) if $\theta_2 < \theta^{b^{-1}}(\theta_1)$, then $\hat{e}_1(\theta) > e^*(\theta)$ and $\hat{e}_2(\theta) < e^*(\theta)$; (ii) if $\theta_2 \in [\theta^{b^{-1}}, \theta^a(\theta_1)]$, then $\hat{e}_1(\theta) < e^*(\theta)$ and $\hat{e}_2(\theta) < e^*(\theta)$; last (iii) if $\theta_2 > \theta^a(\theta_1)$, then $\hat{e}_1(\theta) < e^*(\theta)$ and $\hat{e}_2(\theta) > e^*(\theta)$.

2- Suppose now that $\beta > 1$.

- Under complete information the first best efforts are still symmetric. Moreover $e^*(\beta) = e_1^{FB}(e^*, \beta)$ and $\frac{\partial e^*}{\partial \beta} = \frac{\partial e_1^{FB}}{\partial e_1} \frac{\partial e^*}{\partial \beta} + \frac{\partial e_1^{FB}}{\partial \beta}$ where

$$[\pi_{11}(e_1^{FB}, e_j)V - (1+\lambda)\psi_1(e_1^{FB}, \beta\theta_1 + \theta_2)]\frac{\partial e_1^{FB}}{\partial \beta}$$
$$- (1+\lambda)\psi_{12}(e_1^{FB}, \beta\theta_1 + \theta_2)\theta_1 = 0.$$

Then e_1^{FB} is increasing in β, and $\frac{\partial e^*}{\partial \beta} > 0$.

- Under asymmetric information, the optimal effort are such that

$$[\pi_{11}(e_1^{SB}, e_2)V - (1+\lambda)\psi_{11}(e_1^{SB}, m)]\frac{\partial e_1^{SB}}{\partial \beta}$$
$$- (1+\lambda)\psi_{12}(e_1^{SB}, m)\theta_1 + \lambda\psi_{12}(e_1^{SB}(e_2, \theta), m)\frac{1-F(\theta_1)}{f(\theta_1)} = 0$$
$$[\pi_{22}(e_1, e_2^{SB})V - (1+\lambda)\psi_1(e_2^{SB}, m)\frac{\partial e_2^{SB}}{\partial \beta} - (1+\lambda)\psi_{12}(e_2^{SB}, m)\theta_1 = 0$$

We have $\frac{\partial e_2^{SB}}{\partial \beta} > 0$ and $\frac{\partial e_1^{SB}}{\partial \beta} < \frac{\partial e_2^{SB}}{\partial \beta}$.

$$\frac{\partial \hat{e}_2}{\partial \beta}\left[1 - \frac{\partial e_2^{SB}}{\partial e_1}\Big|_{\hat{e}_1} \cdot \frac{\partial e_1^{SB}}{\partial e_2}\Big|_{\hat{e}_2}\cdot\right] = \frac{de_2^{SB}}{de_1}\Big|_{\hat{e}_1} \cdot \frac{\partial e_1^{SB}}{\partial \beta} + \frac{\partial e_2^{SB}}{\partial \beta} > 0,$$

$$\frac{\partial e^*}{\partial \beta}\left[1 - \frac{\partial e_2^{FB}}{\partial e_1}\Big|_{e^*} \cdot \frac{\partial e_1^{FB}}{\partial e_2}\Big|_{e^*}\cdot\right] = \frac{de_2^{FB}}{de_1}\Big|_{e^*} \cdot \frac{\partial e_1^{FB}}{\partial \beta} + \frac{\partial e_2^{FB}}{\partial \beta}.$$

Since $\frac{\partial e_2^{SB}}{\partial \beta} = \frac{\partial e_2^{FB}}{\partial \beta} = \frac{\partial e_1^{SB}}{\partial \beta}$, we get that $\frac{\partial \hat{e}_2}{\partial \beta} > \frac{\partial e^*}{\partial \beta}$. There exists $\overline{\beta}$ such that if $\beta > \overline{\beta}$, then $\hat{e}_2(\beta) > e^*(\beta)$. In that case $\hat{e}_1(\beta) = e_1^{SB}(\hat{e}_2(\beta)) < e_1^{SB}(e^*(\beta)) < e_1^{FB}(e^*(\beta)) = e^*(\beta)$ □.