Monitoring with Collusion and the Value of Information

Walter Cont
FIEL and UNLP
e-mail: walter@fiel.org.ar

Abstract

In a principal-supervisor-agent relationship with collusion, I rank the principal's preference over the quality of the supervisor's information. In the presence of imperfect signal and output distortions, the principal does better with hard-but-non-forgeable information than with soft information. Only in the limit case of accurate signal (i.e., the supervisor may observe either the true type or nothing) the principal with soft information is as well off as with hard information even under output distortions (as in Baliga). Nevertheless, distortions are needed for the creation of collusion stakes through differential information rents when the supervisor's signal is noisy.

The conditions under which the supervisor with soft information is still valuable for the principal are, first, that the supervisor's signal's must exceed some lower treshold of noise and, second, that side transfers between the agent and the supervisor must be inefficient.

1 Introduction

In many economic activities there exists the possibility of collusion. Examples abound: supervisors / auditors collude with employees in detriment of company owners, a regulatory agency can be captured by a regulated company or specific interest groups, a ticket inspector can be bribed by a free-rider caught, and so on.

In the contract theory literature, it was Tirole [9, 10] who first showed that a supervisor is valuable to the principal when the supervisor's information about the worker's productivity is "hard" (in other words, when the supervisor can hide verifiable evidence) even if there is the possibility of collusion between the supervisor and the agent.

Baliga [2] used Tirole [10] toy model to show that the supervisor is useful to the principal even though information is "soft" (in other words, when the information provided by the supervisor cannot be backed up) and both agent and supervisor can collude. This result contrasted Tirole's conjecture that "... If the supervisor and the agent collude, however, soft information becomes useless,..." (Tirole [9], p. 190).

Tirole [10] model has two main characteristics I discuss here. First, it is a standard adverse selection model. The agent supplies 0 or 1 unit of a good, having private information of its productivity. Second, the supervisor's information structure enables him to learn the right information or nothing.

In this paper I analyze the value of soft information to in a more general framework. In particular, I modify the assumptions on production technology and supervisor's information structure. On the one hand, I allow the production cost (or profit) to depend on both the agent's productivity and effort. In particular, differently from Tirole [10] and Baliga [2], effort can be used to increase or decrease profits (i.e., distort allocations). Second, I introduce imperfect signals: the supervisor may learn the information correctly (as in these two papers) or incorrectly (see Kofman and Lawarrée [8]), or may learn nothing.

In order to present the results of this paper, it is helpful to rename the quality of information, allowing for three different informational environments:

(i) "hard and non-forgeable" information: the supervisor has verifiable evidence of his signal and may conceal it from the principal (see Tirole [9, 10])

(ii) "hard and forgeable" information: the monitor can falsify his signal with help from the agent (see Kofman and Lawarrée [8])

(iii) "soft" information: the monitor has no verifiable proof of his signal, and hence may report anything (see Baliga [2], Faure-Grimaud *et al.* [6]).

This clasification of information structures implies that the principal may obtain a different level of utility depending on the quality of the information that the supervisor reports to her. Specifically, the third case is a more serious problem to the principal than the second case is, which is more serious than the first. The quality of the monitor's information is lower when it becomes "less reliable". Consequently, with collusion, the cost to the principal of obtaining a truthful report from the monitor is higher. Nevertheless, it is important to study each of these cases individually. Depending on the activity to be controlled, the monitor may "ignore" relevant information to write his report (such as not reporting perquisites), or may not obtain information when it is hard to dispose of.[1] These are examples of hard and non-forgeable information. In addition, the monitor may write reports based on evidence pre-selected by the manager (such as audit reports in a company's credit department), or distort information (alter payrolls, create fictitious personnel, manipulate quality tests, etc., see Dalton [5], p. 32). These last two cases correspond to hard and forgeable information and soft information, respectively.

I find that there is a clear preference order by the principal of the information structures when profits or costs can be affected by the agent's effort *and* the supervisor's information is noisy (in the sense that he may make mistakes): hard and non-forgeable information is preferred to hard and forgeable information.

Second, I show an ***equivalence result*** between hard and forgeable information and soft information. This result contrasts both Tirole's conjecture and Baliga's result. Only in particular circumstances soft information makes the supervisor valueless to the principal, those of (i) perfect side-contracting technology (ii) imperfect side-contracting technology and very noisy signal. Also, soft information is equivalent to hard (and non-forgeable) information when the supervisor observes the agent's type without error (although he can observe nothing).

The paper is organized as follows. Section 2 outlines the model. Section 3 sets the benchmarks (No Monitor and No Collusion). I study the effect of collusion under different information structures (hard and non-forgeable, hard and forgeable and soft information) in Sections 4 to 6. Finally, Section 7 concludes.

[1] Dalton [5] described that "...safety and health inspectors usually telephoned in advance of visits so that they would not see unsafe practices or conditions they would feel obliged to report." (p. 48).

2 Model

2.1 Technology

Consider a hierarchy consisting of three parties: a principal, a supervisor and an agent. The principal hires the agent to produce a unit of a good. The agent combines productivity/expertise θ and effort/investments e to get a gross profit $x = \theta e$.

The principal observes profit x but does not observe productivity nor effort, so she can pay a net transfer t to the agent based only on performance x.

The principal knows that θ can take either of two values θ_H or θ_L, with $\theta_H > \theta_L > 0$, and has a prior distribution characterized by $q = \Pr\ (\theta = \theta_H)$, with $q \in (0, 1)$. The parameter θ_H represents a higher productivity. If the agent exerts effort $e \geq 0$, it produces gross profit θe but derives a private disutility or cost $\psi(e) = e^2/2$.

The supervisor is introduced as a third agent hired by the principal to obtain information about the agent's productivity. The supervisor obtains a *noisy signal* of the agent's productivity (this signal is also observed by the agent) at no cost.[2] I assume that the supervisor's information is valuable to the principal because it reduces the information assymetry between her and the agent. The supervisor's signal may take one of the following values: With probability $1-p$ the monitor learns nothing about the agent's type (i.e., $\sigma = \emptyset$). Given the agent's type, the supervisor observes the true type with probability $p\alpha$ (i.e., $\sigma = \theta$) and the incorrect type with probability $p(1-\alpha)$ (e.g., if $\theta = \theta_H$, $\sigma = \theta_L$), where $\alpha > 1/2$.[3] As $p\alpha \to 1$ the quality of the signal is better, in that the supervisor "learns" the right agent's type. The supervisor sends a report $r \in \{\emptyset, \theta_L, \theta_H\}$ to the principal, who pays him a compensation $w \geq 0$ (the supervisor is protected by limited liability).

2.2 Collusion and Information

As shown below, the supervisor's report affects the agent's rents. I allow for the possibility of collusion between the agent and the supervisor. They may coordinate to manipulate a

[2] This asumption corresponds to the case of a principal hiring a supervisor for other reasons than supervision (e.g., coordination, advising, etc.). The results extend to a costly supervisor, provided that he is hired.

[3] This assumption satisfies the monotone likelihood ratio property that a correct signal is more probable.

report in order to appropriate these rents. The way this report can be changed depends on the structure of the supervisor's information and the enforceability of side-contracts.[4]

I analyze three possible information structures. First, the supervisor's signal may be hard and not forgeable. In this case, the signal is *verifiable* to the principal. The supervisor may conceal information and report that he observed "nothing", but cannot forge it and report that he observed other type.[5]Second, information may be hard and forgeable, in which case the supervisor may make up a verifiable report with the agent's help. Third, information may be soft, i.e., the supervisor cannot back up his report with a verifiable proof.

Side transfers may be *inefficient*. A transfer b from the agent to the supervisor is valued $v(b) = kb$ by the latter, where $0 \leq k \leq 1$.

2.3 Timing and Utilities

The timing of the game is as follows. At time $\tau = 0$, the agent learns θ and σ and the supervisor learns σ. At $\tau = 1$ the principal offers individual contracts to the agent and the supervisor. The three parties sign the contract. At $\tau = 2$ the agent chooses effort e and produces a gross profit x. Transfers –and side transfers– are made.

All parties are risk neutral. The principal, agent and monitor's utility is $U_P = x - t - w$, $U_A = t - b - e^2/2$ and $U_M = w + kb$, respectively. The agent and supervisor's reservation utilities are $\underline{U}_A = \underline{U}_S = 0$, respectively.

2.4 First Best

When the principal observes both agent's effort and type, the problem simplifies to choose effort and transfers in order to maximize $\theta_j e - t$ subject to the agent's participation constraint $t - e^2/2 \geq 0$, for $j = L, H$. The solution to this problem is: $e_j^{FB} = \theta_j$, $t_j^{FB} = \theta_j^2/2$, for $j = L, H$. The principal's expected utility is $EU_P^{FB} = \left[q\theta_H^2 + (1-q)\theta_L^2\right]/2$.

[4] In order to focus on the effects of collusion on contract design, I assume that side-contracts are enforceable (although transfers may be inefficient) and non-renegotiable, and hence I obtain an upper (lower) bound on the coalition's (principal's) utility. For a discussion on this point, see Tirole (1992).

[5] The possibility that the supervisor threatens the agent with reporting that he observed nothing when the agent earns high rents, so the agent pays the supervisor to report the right signal, is ruled out.

3 Benchmarks: No Supervisor, No Collusion

In this section I present two relevant benchmarks. The lower bound on the principal's utility is achieved by contracting with the agent directly. The upper bound is achieved when the agent and the supervisor cannot collude, for example, because the supervisor is honest. In this case, the principal obtains the monitor's signal at no cost.

3.1 No Supervisor

When the principal does not observe either agent's effort or type, the contract offered to the agent should be based only on the observable performance x. The solution for this problem is standard. From the Revelation Principle, the principal can restrict herself to direct mechanisms based on an agent's truthful report of her type. For a report $\hat{\theta}_j$, $j = L, H$, there is an effort recommendation $e_j = e_j(\hat{\theta}_j)$ to achieve a gross profif $x_j = \hat{\theta}_j e_j$, and a payment t_j. A feasible contract must satisfy the individual rationality (IR) and incentive compatibility (IC) constraints

$$\begin{array}{llll} IR(L): & t_L \geq e_L^2/2 & IC(L): & t_L - e_L^2/2 \geq t_H - e_H^2/2\Delta\theta \\ IR(H): & t_H \geq e_H^2/2 & IC(H): & t_H - e_H^2/2 \geq t_L - e_L^2\Delta\theta/2 \end{array}$$

The principal maximizes $q(\theta_H e_H - t_H) + (1-q)(\theta_L e_L - t_L)$ subject to these contstraints. The solution to this problem is found by solving the restricted problem with the binding constraints IR(L) and IC(H).[6] The restricted problem simplifies to choose e_L and e_H to maximize

$$q\left[\theta_H e_H - \frac{e_H^2}{2} - R\frac{e_L^2}{2}\right] + (1-q)\left[\theta_L e_L - \frac{e_L^2}{2}\right]$$

and the solution to this problem and the principal's utility are:

$$\left\{ e_L^{NS} = \frac{(1-q)\theta_L}{(1-q)+qR} \ , \ t_L^{NS} = \frac{\left(e_L^{NS}\right)^2}{2} \right\} \quad ; \quad \left\{ e_H^{NS} = \theta_H \ , \ t_H^{NS} = \frac{\theta_H^2}{2} + R\frac{\left(e_L^{NS}\right)^2}{2} \right\}$$

$$EU_P^{NS} = q\frac{\theta_H^2}{2} + \frac{(1-q)^2\theta_L^2}{2\left((1-q)+qR\right)} \tag{1}$$

where the superscript NS stands for "No Supervisor". The intuition for this result is simple. In order to elicit first-best effort from the more productive agent (who has incentives to claim that she is inefficient), the principal pays her an information rent. This rent

[6] The proof that constraints IC(L) and IR(H) are slack is standard and hence omitted.

depends on the agent's utility of accepting the contract designed to the less productive agent (specifically, from IC(H) the information rent $R_H = t_H - e_H^2/2$ is equal to $Re_L^2/2$). By eliciting lower effort and forgoing profit from the less productive agent the principal makes this contract less attractive to the more productive agent, and hence pays her a lower information rent. The optimal contract without supervisor resolves this trade-off.

3.2 Honest Supervisor: No Collusion

Hiring a supervisor helps the principal to achieve a higher utility than EU_P^{NS} from Equation (1). This section solves the optimal contract with an honest supervisor shows the maximum utility that the principal can achieve from this contract.

The supervisor reports his signal $r = \sigma$ truthfully. The problem is *as if* the principal was vested with the new information structure: six states of the world which correspond to the combination of two types and three signals. Therefore, the problem and its solution is a straight extension of those in Section 3.1.

Let e_{jr} be the agent's effort and t_{jr} be her compensation when the agent reports $\hat{\theta}_j \in \{\theta_L, \theta_H\}$ and the supervisor reports $r \in \{\emptyset, \theta_L, \theta_H\}$ in a direct mechanism. Similarly, the supervisor's compensation is w_{jr}. The agent's participation and incentive constraints in a feasible contract are summarized by

$$\begin{array}{ll} \text{IR(jr)}: & t_{jr} \geq e_{jr}^2/2 \\ \text{IC(Hr)}: & t_{Hr} - e_{Hr}^2/2 \geq t_{Lr} - e_{Lr}^2 \Delta\theta/2 \end{array} \quad j = L, H, \;\; r = \emptyset, L, H \tag{2}$$

and the supervisor's participation and limited liability constraints are summarized by

$$w_{jr} \geq 0,^{7} \qquad j = L, H, \quad r = \emptyset, L, H \tag{3}$$

Let π_{jr} denote the probability of occurrence of each state, where $\pi_{L\emptyset} = (1-q)(1-p)$, $\pi_{LL} = (1-q)p\alpha$, $\pi_{LH} = (1-q)p(1-\alpha)$, $\pi_{H\emptyset} = q(1-p)$, $\pi_{HL} = qp(1-\alpha)$, and $\pi_{HH} = qp\alpha$. The principal maximizes $\sum_j \sum_r \pi_{jr} \left(\theta_j e_{jr} - t_{jr} - w_{jr}\right)$ subject to these contstraints.

The solution of this problem is a straightforward extension of that in Section 3.1, and is found solving a restricted problem with binding constraints IR(Lr) and IC(Hr).[8] The

[7] The participation constraint is $qw_{Hr} + (1-q)w_{Lr} \geq 0$ for every signal (because the supervisor does not know the agent's type). These constraints are satisfied with limited liability ($w \geq 0$).

[8] When constraints IR(Lr) and IC(Hr) are binding, constraints IR(Hr) and IC(Lr) are slack, for $r = \emptyset, L, R$. This proof is standard and hence omitted.

principal pays $w_{jr}^{NC} = 0$ to the suprevisor. The restricted problem reduces to choose $\{e_{jr}\}$ to maximize

$$\sum_{r\in\{0,L,H\}} \pi_{Hr}\left[\theta_H e_{Hr} - \frac{e_{Hr}^2}{2} - R\frac{e_{Lr}^2}{2}\right] + \sum_{r\in\{0,L,H\}} \pi_{Lr}\left[\theta_L e_{Lr} - \frac{e_{Lr}^2}{2}\right]$$

In the solution to the optimal contract with an honest supervisor, the principal elicits first-best effort from the more productive agent for all supervisor's signals.

$$e_{H\emptyset}^{NC} = e_{HL}^{NC} = e_{HH}^{NC} = \theta_H \tag{4}$$

while she recommends an effort to the less productive agent according to

$$e_{L\emptyset}^{NC} = \frac{(1-q)\theta_L}{(1-q)+qR}; \quad e_{LL}^{NC} = \frac{(1-q)\alpha\theta_L}{(1-q)\alpha + q(1-\alpha)R}, \quad e_{LH}^{NC} = \frac{(1-q)(1-\alpha)\theta_L}{(1-q)(1-\alpha)+q\alpha R} \tag{5}$$

On the other hand, the principal leaves no rent to the less productive agent and pays an information rent to the more productive agent

$$t_{Lr}^{NC} = \frac{\left(e_{Lr}^{NC}\right)^2}{2}, \quad t_{Hr}^{NC} = \frac{\theta_H^2}{2} + R_{Hr}^{NC}, \qquad \text{for } r = \emptyset, L, H \tag{6}$$

where $R_{Hr}^{NC} = R\left(e_{Lr}^{NC}\right)^2/2$. The principal's utility is

$$EU_P^{NC} = q\frac{\theta_H^2}{2} + \frac{(1-q)^2\theta_L^2}{2}\left\{\frac{1-p}{(1-q)+qR} + \frac{p\alpha^2}{(1-q)\alpha+q(1-\alpha)R} + \frac{p(1-\alpha)^2}{(1-q)(1-\alpha)+q\alpha R}\right\} \tag{7}$$

Next Proposition summarizes the utility achieved by the principal in this no-collusion environment (see Tirole [9]).

Proposition 1 *The optimal contract when the principal hires an honest supervisor is such that* $w_{jr}^{NC} = 0$ *and* $\{e_{jr}^{NC}, t_{jr}^{NC}\}$ *satisfies (4)-(6), for* $j = L, H$ *and* $r = \emptyset, L, H$.

As in the no-supervisor case, the principal solves a trade-off between paying a lower rent to the more productive agent and distorting effort, and forgoing profit, to the less productive agent, *for every supervisor's report*. Equations (5) and (6) show the solution to this trade-off, as the principal elicits $e_{LL} > e_{L\emptyset} > e_{LH}$ from the less productive agent and pays $R_{HL} > R_{H\emptyset} > R_{HH}$ to the more productive agent.

The intuition of this result is as follows. The principal updates probabilities of facing a more or less productive agent. Suppose that the supervisor observes $\sigma = \theta_H$, and recall that the more productive agent's degree of discretion is to misreport state HH stating that it is LH. The new infomration structure allows the principal to infer that the probability of state LH is less than $1 - q$ and that the probability of state HH is greater than q. Hence she distors effort e_{LH} the low-probability state LH (more than in the no-monitor case), which allows her to pay lower rents R_{HH} in the high-probability state HH. Suppose now that the supervisor observes $\sigma = \theta_L$. The principal finds it profitable to create less effort distortion in the high-probability state LL (she increases e_{LL}), at the cost of paying higher information rents R_{HL} in the low-probability state HL. If the supervisor observes the right type or nothing ($\alpha = 1$, and $p > 0$), the high inefficiencies and rents are ex ante costless (states LH and HL do not occur). Next Corollary summarizes the result that supervision is valuable for the principal.

Corollary 1 *The principal hires a costless honest supervisor always.* $EU_P^{NC} > EU_P^{NS}$ *for* $\alpha > 1/2$ *and* $p > 0$.

Corollary 1 states that as long as the supervisor's signal is informative enough (i.e., it observes something –$p > 0$– and the signal has some value –$\alpha > 1/2$–) the supervisor is always valuable to the principal. If there is a cost c to hire the supervisor, then the principal hires the latter if $EU_P^{NC} - EU_P^{NS} > c$. The supervisor's signal must surpass a reliability treshold ($p\alpha$ must exceed some lower bound). Finally, the principal's utility approaches the first-best EU_P^{FB} as the probability of observing the agent's type increases ($p\alpha \to 1$).

4 Collusion: Hard and Non-Forgeable Information

In the remaining of the paper I allow for collusion between the agent and the supervisor, and look for collusion-proof optimal contracts (which in fact are optimal contracts). In this Section I assume that the supervisor's information is hard and non-forgeable, which means that the supervisor discretion lies in concealing his signal from the principal. I also assume that the supervisor has full bargaining power and makes take-it-or-leave-it offers when he bargains with the agent to conceal information.

A feasible contract with collusion and hard-but-non-forgeable information must satisfy conditons (2) and (3), namely, individual incentive, participation and liability constraints. In addition, given that the supervisor's discretion lies in concealing his signal, a feasible collusion-proof contract must compensate the agent-supervisor coalition so that concealment of the signal is not profitable for the coalition members. These coalition constraints are

$$\begin{aligned}&Pr(\theta_H|\sigma = H)\,[w_{HH} + kR_{HH}] + Pr(\theta_L|\sigma = H)\,[w_{LH} + kR_{LH}] \geq \\ &Pr(\theta_H|\sigma = H)\,[w_{H\emptyset} + kR_{H\emptyset}] + Pr(\theta_L|\sigma = H)\,[w_{L\emptyset} + kR_{L\emptyset}]\end{aligned} \tag{8}$$

$$\begin{aligned}&Pr(\theta_H|\sigma = L)\,[w_{HL} + kR_{HL}] + Pr(\theta_L|\sigma = L)\,[w_{LL} + kR_{LL}] \geq \\ &Pr(\theta_H|\sigma = L)\,[w_{H\emptyset} + kR_{H\emptyset}] + Pr(\theta_L|\sigma = L)\,[w_{L\emptyset} + kR_{L\emptyset}]\end{aligned} \tag{9}$$

because the supervisor does not know the agent's type at the side-contract stage and assesses probabilities of facing each type θ conditioned on having observed a signal σ. The optimal contract solves

$$\max_{\{e_{jr}, t_{jr}, w_{jr}\}} \sum_{j\in\{L,H\}} \sum_{r\in\{\emptyset,L,H\}} \pi_{jr} \{\theta_j e_{jr} - t_{jr} - w_{jr}\}$$

subject to (2), (3), (8) and (9). Define α_h such that the coalition constraint (8) is binding with $w_{HH} > 0$ for $\alpha > \alpha_h$, where

$$\alpha_h = \begin{cases} \frac{-[2(1-p)-k]+\sqrt{\left[2(1-p)-k\right]^2+4pk(1-p)}}{2pk} & \text{if} \quad k > 0 \\ 1/2 & \text{if} \quad k = 0 \end{cases} \tag{10}$$

The following Proposition summarizes the optimal contract with collusion and hard but non-forgeable information (Tirole (1988)).

Proposition 2 *The optimal contract with hard but non-forgeable information satisfies Proposition 1 for $p = 1$. Otherwise, let $r = \emptyset, L, H$ be the index associated to the supervisor's report. The principal*

- *elicits first-best effort from the more productive agent $e^{NFI}_{H\emptyset} = e^{NFI}_{HL} = e^{NFI}_{HH} = \theta_H$ and pays her an information rent $R^{NFI}_{Hr} = R\left(e^{NFI}_{Lr}\right)^2/2$,*
- *pays the effort cost $t^{NFI}_{Lr} = \left(e^{NFI}_{Lr}\right)^2/2$ to the less productive agent,*

- *pays* $w_{Lr}^{NFI} = w_{H\emptyset}^{NFI} = w_{HL}^{NFI} = 0$ *to the supervisor (i.e., in all states but HH),*
- *distorts the less productive agent's effort and pays a wage* w_{HH}^{NFI} *to the supervisor depending on* α*, as follows,*

  * *for* $\alpha \leq \alpha_h$, $e_{LL}^{NFI} > e_{L\emptyset}^{NFI} = e_{LH}^{NFI}$ *and* $w_{HH}^{NFI} = 0$*, where*

$$
\begin{aligned}
e_{LL}^{NFI} &= \frac{(1-q)\alpha\theta_L}{(1-q)\alpha + q(1-\alpha)R} \\
e_{L\emptyset}^{NFI} = e_{LH}^{NFI} &= \frac{(1-q)(1-p\alpha)\theta_L}{(1-q)(1-p\alpha) + qR(1-p(1-\alpha))}
\end{aligned}
$$

  * *for* $\alpha > \alpha_h$, $e_{LL}^{NFI} > e_{L\emptyset}^{NFI} > e_{LH}^{NFI}$ *and* $w_{HH}^{NFI} = kR\frac{\left(e_{L\emptyset}^{NFI}\right)^2 - \left(e_{LH}^{NFI}\right)^2}{2}$*, where*

$$
\begin{aligned}
e_{LL}^{NFI} &= \frac{(1-q)\alpha\theta_L}{(1-q)\alpha + q(1-\alpha)R} \\
e_{L\emptyset}^{NFI} &= \frac{(1-q)(1-p)\theta_L}{(1-q)(1-p) + qR\left(1-p(1-\alpha k)\right)} \\
e_{LH}^{NFI} &= \frac{(1-q)(1-\alpha)\theta_L}{(1-q)(1-\alpha) + q\alpha R(1-k)}
\end{aligned}
$$

  * *in particular,* $e_{LL}^{NFI} = e_{LL}^{NC}$ *and* $R_{HL}^{NFI} = R_{HL}^{NC}$*.*

Proof: See Appendix.

Consider the contract without collusion as a benchmark to analyze the design of the optimal contract when the agent and the supervisor can conceal the supervisor's information. The more productive agent's rent with signal $\sigma = L$ is higher than her rent with signal $\sigma = \emptyset$, that is, $R_{HL} > R_{H\emptyset}$. In this case, there are no incentives to conceal information. The principal does not need to change e_{LL} and R_{HL} from those in the no-collusion contract.[9]

[9] This result depends on the assumption that no blackmail is possible. If the supervisor could threaten the more productive agent with concealing his signal $\sigma = L$, he could expect the agent to bribe him up to the difference of information rents in states HL and H$\emptyset$. Given the assumption that the supervisor has full bargaining power, there is no incentive for the agent to bribe the supervisor in order to induce him to report the right signal. But the principal wants the supervisor to report the right signal. In the optimal contract the principal pays a rent $R_{H\emptyset}$ to the agent and a wage $w_{HL} = k(R_{HL} - R_{H\emptyset})$ to the supervisor in state HL (everything else is as in Proposition 2). Moreover, the principal derives higher utility than that from Proposition 2 for $k < 1$.

In the collusion-free contract, the more productive agent's information rent with signal $\sigma = H$ is lower than her rent with signal $\sigma = \emptyset$, that is, $R_{HH} < R_{H\emptyset}$. In this case, there are incentives to conceal information. The principal may overcome collusion and obtain the true signal in two different ways. On the one hand, she can pay the supervisor the agent's *differential* rent between states H∅ and HH, adjusted by the inefficiency in side transfers: $w_{HH} = k(R_{H\emptyset} - R_{HH})$. In this case, the aggregate information cost in state HH is the "coalition information rent" $R_{HH} + w_{HH} = R_{HH} + k(R_{H\emptyset} - R_{HH})$. The intuition of this cost is easy to grasp. The threat of collusion forces the principal to spread the type-θ_H agent's rent from state H∅ over state HH (up to the inefficiency in side transfers), which turns out to be $R_{H\emptyset}$ when side transfers are efficient. On the other hand, the principal may eliminate the incentives to collude, by recommending the less productive agent to exert the same effort in states L∅ and LH, i.e., $e_{L\emptyset} = e_{LH}$. The first option is profitable when the supervisor's signal is relatively accurate ($\alpha > \alpha_h$), and the second alternative is profitable when distorting efforts is –ex ante– too costly to pay the supervisor the additional rent, which happens when his signal is not accurate enough ($\alpha \leq \alpha_h$). Nevertheless, in either case, the principal finds profitable to hire the supervisor.

From equation (10), the cut-off value α_h increases in k (in particular, $\alpha_{h(k=0)} = 1/2$, and $\alpha_{h(k=1)} = 1$). For $\alpha \leq \alpha_h$ the principal does not change allocations as k increases, while for $\alpha > \alpha_h$ the principal either pays a higher coalition rent or switchs not to distort allocations. Therefore, the principal's utility is non-increasing in k. Moreover, the principal derives less utility under collusion than under no collusion for any $p < 1$.[10] This result is summarized next.

Corollary 2 *When the supervisor's information is hard but non-forgeable, as long as information is not perfectly informative of the agent type ($p\alpha < 1$),*

(i) the principal derives less utility with the collusion-proof contract than with the collusion-free contract, and

(ii) the principal's utility is non-increasing in the side-transfer efficiency parameter k.

[10] This can be verified by comparing allocations and compensations under both problems (or else compare (19) with (7)).

5 Hard and Forgeable Information

Now I turn to the case in which the agent-supervisor coalition can manipulate the supervisor's signal in any direction. They have incentives to distort the supervisor's report whenever there is a state of the world in which the agent-supervisor coalition's rent is higher.

Given conditions (8)-(9) and the fact that in Sections 3.2 and 4 the principal paid rents $R_{HL} > R_{H\emptyset} \geq R_{HH}$, where R_{Hr} is the agent's rent in state Hr, for $r \in \{\emptyset, L, H\}$, it is straightforward to figure out that the coalition will have incentives to distort the supervisor's report $\emptyset$ or H to L when the agent is more productive (i.e., the supervisor provides incorrect information about the agent's type). In addition, it was shown in the Appendix that, after simplifications, the collusion-proof contract with constraints (8)-(9) can be found by solving the restricted program with constraints (16)-(17) (which resemble those *as if* the supervisor knew the agent's type). These two results lead to the following simplified coalition constraints in the collusion proof-contract with hard and forgeable information:

$$w_{H\emptyset} + kR_{H\emptyset} \geq w_{HL} + R_{HL} \tag{11}$$

$$w_{HH} + kR_{HH} \geq w_{HL} + R_{HL} \tag{12}$$

Define α_f^1 such that constraint (12) is binding with $w_{HH} > 0$ for $\alpha > \alpha_f^1$, and α_f^2 such that constraint (11) is binding with $w_{H\emptyset} > 0$ for $\alpha > \alpha_f^2$ (Equations (21) and (20) in the Appendix, respectively):

$$\alpha_f^1 = \frac{1}{2-k} \qquad\qquad \alpha_f^2 = \frac{k + (1-k)p}{2p(1-k)} \tag{13}$$

Next Proposition summarizes the optimal contract with collusion and hard-and-forgeable information.

Proposition 3 *In a principal-supervisor-agent relationship with collusion and hard-and-forgeable information,*

- *if $\alpha \leq \alpha_f^1$, the principal does not hire the supervisor;*

- *if $\alpha > \alpha_f^1$, the principal induces the type-θ_H agent to exert effort $e_{H\emptyset}^{FI} = e_{HL}^{FI} = e_{HH}^{FI} = \theta_H$, pays her an information rent $R_{Hr}^{FI} = R\left(e_{Lr}^{FI}\right)^2/2$ (which depends on both* *the supervisor's report $r = \emptyset, L, H$, and the type-θ_L agent's effort), pays no rent to the type-θ_L agent, $t_{Lr}^{FI} = \left(e_{Lr}^{FI}\right)^2/2$, pays $w_{L\emptyset}^{FI} = w_{LL}^{FI} = w_{LH}^{FI} = w_{HL}^{FI} = 0$ to the supervisor (i.e., in all states but $H\emptyset$ and HH), and distorts the type-θ_L agent's effort and the supervisor's wage depending on the value ofα:*

- *when $\alpha \leq \alpha_f^2$,*

$$\begin{aligned}
e_{L\emptyset}^{FI} &= e_{LL}^{FI} = \frac{(1-q)(1-p(1-\alpha))\theta_L}{(1-q)(1-p(1-\alpha)) + qR\left[1-p\alpha(1-k)\right]} \\
e_{LL}^{FI} &= \frac{(1-q)p\alpha\theta_L}{(1-q)p\alpha + qR\left[k+p(1-\alpha)(1-k)\right]} \\
w_{H\emptyset}^{FI} &= w_{HH}^{FI} = k(R_{HL}^{FI} - R_{HH}^{FI})
\end{aligned}$$

- *when $\alpha > \alpha_f^2$,*

$$\begin{aligned}
e_{L\emptyset}^{FI} &= \frac{(1-q)\theta_L}{(1-q) + qR(1-k)} \\
e_{LL}^{FI} &= \frac{(1-q)p\alpha\theta_L}{(1-q)p\alpha + qR\left[k+p(1-\alpha)(1-k)\right]} \\
e_{LH}^{FI} &= \frac{(1-q)(1-\alpha)\theta_L}{(1-q)(1-\alpha) + q\alpha R(1-k)} \\
w_{H\emptyset}^{FI} &= k(R_{HL}^{FI} - R_{H\emptyset}^{FI}) \\
w_{HH}^{FI} &= k(R_{HL}^{FI} - R_{HH}^{FI})
\end{aligned}$$

Proof: See Appendix.

First, let us interpret the coalition constraints. Suppose that a contract in state H$\emptyset$ is such that $R_{H\emptyset}^{FI} < R_{HL}^{FI}$ and $w_{H\emptyset}^{FI} = w_{HL}^{FI}$. The coalition will change the supervisor's report to $r = L$. On the other hand, if $R_{H\emptyset}^{FI} > R_{HL}^{FI}$ and and $w_{H\emptyset}^{FI} = w_{HL}^{FI}$, the coalition will change the report to $r = \emptyset$. When the principal avoids collusion in one direction, she creates stakes in collusion in the other direction. This result holds for every supervisor's signal when the agent's type is θ_H. Therefore, an optimal contract must satisfy the constraint

$$w_{H\emptyset} + kR_{H\emptyset} = w_{HL} + kR_{HL} = w_{HH} + kR_{HH}$$

This contract, which is designed to avoid changing the supervisor's report to $r = L$, also prevents coalition deviations in any direction (e.g., in state HH there will not be incentives to claim that the state is H$\emptyset$, etc.).

Proposition 3 states that when information can be forged, the accuracy of the supervisor's signal must exceed some minimum level α_f^1 for the supervisor to be valuable to the principal. Two reasons explain this result. First, a noisy signal is not very informative of the agent's type. Second, when the supervisor's report can be manipulated, the informational value of the signal is reduced.

The contract is more flexible (in the sense that effort and rents are adjusted to posterior beliefs) as the signal becomes more informative of the agent's type. The principal only distorts effort e_{LH} (from $e_{L\emptyset} = e_{LL}$) for $\alpha_1^h < \alpha \leq \alpha_2^h$, while she sets $e_{LL} > e_{L\emptyset} > e_{LH}$ when $\alpha > \alpha_2^h$.

The principal finds profitable to hire the supervisor if side transfers are inefficient. In the limiting case of efficient side transfers ($k = 1$),the principal has to transfer the whole agent's rent to the supervisor to avoid forgery, and hence deals with the agent directly (note that $\alpha_1^f(k = 1) = 1$).[11]

Finally, when $\alpha = 1$, forgery is not possible (but concealment is) because the supervisor, who observes the agent's type correctly, cannot claim having seen $\sigma = L$ ($\sigma = H$) when the agent reports θ_H (θ_L) truthfully, but may claim having observed $\sigma = 0$. In this case, the contract is that of Proposition 2.

From equation (21), the cut-off value α_h^1 increases in k (in particular, $\alpha^1_{f(k=0)} = 1/2$, and $\alpha^1_{f(k=1)} = 1$). For $\alpha \leq \alpha_f^1$ the principal does not change allocations as k increases, while for $\alpha > \alpha_f^1$ the principal either pays a higher coalition rent or switchs not to hire the supervisor. Moreover, for $\alpha > \alpha_f^2$ the principal either pays a higher coalition rent or switchs not to distort allocations. These results imply that the principal's utility is non-increasing in k. Moreover, the principal derives less utility under hard and forgeable information than under hard and non-forgeable information for $\alpha < 1$.[12] This result is summarized next.

Corollary 3 *When information is hard and forgeable*

[11] In general, for a given inefficiency of side transfers, this result holds if the supervisor's signal is noisy enough (i.e., $\alpha < \alpha_1^f$).

[12] This can be verified by comparing allocations and compensations under both problems (or else compare (22) with (19)).

(i) if $\alpha < 1$, the principal derives less utility with hard and forgeable information than with hard and non-forgeable information,

(ii) the principal's utility is non-increasing in the side-transfer efficiency parameter k,

(iii) moreover, if side transfers are efficient ($k = 1$) the supervisor adds no additional value to agency the relationship.

6 Soft Information

In this Section I study whether the fact of information being soft (i.e., the supervisor has no verifiable proof of his signal) reduces the value to the principal of the reports sent by the supervisor. In particular, I am interested in finding the level of utility the principal can achieve, since at first sight she is expected to be more constrained in the choice of contracts.

Taking as reference the collusion-free contract (Section 3.2), two main problems arise with soft information. First, as in Section 5, the agent-supervisor coalition will change reports when the parties find it profitable (as with hard and forgeable information). Second, the supervisor will change a report unilaterally when there are gains from doing so.

Fortunately, next Theorem states that the first problem is relevant in the design of a contract for the agent and supervisor. Specifically, the principal can achieve the hard-and-forgeable-information utility (from Section 5) in a Perfect Bayesian Equilibrium if the signal is noisy ($\alpha < 1$), and she can achieve the hard-and-non-forgeable-information utility if the signal is obtained without noise ($\alpha = 1$). Hence, the results in Section 5 applies to soft information.

Theorem 1 *When the supervisor's signal is soft (i.e., he has no verifiable proof of it),*

- *if the supervisor's signal is noisy ($\alpha < 1$), the principal achieves the same utility as that when information is hard and forgeable;*
- *if the signal is accurate ($\alpha = 1$), the principal achieves the sale utility as when information is hard but non-forgeable.*

Proof: See Appendix.

This ***utility equivalence*** result derives from the fact that the principal can control collusion at no additional cost (compared to the hard-and-forgeable-infromation case). Suppose that the principal offers a collusion-free contract (i.e., a contract that does not take collusion constraints into consideration). Then, the supervisor-agent coalition has incentives to make up a profitable signal (and the agent provides the supervisor with the relevant information to do it).[13] Even easier for the coalition parties, when information is soft, they will coordinate reports without any need to proof them. Moreover, since the agent knows the supervisor's signal, the principal can use an agent's report of that signal in the mechanism. By minimizing compensations when the reports on the supervisor's signal differ, the principal can eliminate unilateral changes in reports. Note that whether information is verifiable or not, the supervisor is valuable to the principal, who pays him accordingly to obtain a truthful report.

This paper shows that only in special cases the equivalence result extends to hard and non-forgeable information (Baliga [2]), in particular when there is no room for mistakes in the information held by the supervisor ($\alpha = 1$). Even in the case of cost distortions the equivalence holds.[14] But when production cost changes with effort *and* the information held by the supervisor is noisy, soft information produces less value than hard and non-forgeable information. This is because distortions in (less productive agent's) efforts create distortions in (more productive agent's) rents. This effect is not present when effort is absent. The possibility of forging information by the coalition parties to get these rents reduces the value of the supervisor's signal under soft information.

But insofar as the side-transfer technology is inefficient ($k < 1$) and the signal is not excessively noisy (α greater than some lower bound) the signal is valuable to the principal (i.e., she earns more profit if she hires the supervisor than if she does not). Finally, this results goes along the lines of Faure-Grimaud *et al.* [7], in the sense that supervision is valuable for organization, but it departs from their work in that I compare different information structures (i.e., hard and non-forgeable, hard and forgeable and soft information) and the authors compare different organizational forms (centralization vs. decentralization).

---

[13] Of course, this result depends on the assumption that the agent provides this information costlessly.

[14] Recall that Tirole's (1992) and Baliga's (1999) models assume adverse selection and no cost-reducing effort. The supervisor observes either the agent's true type or nothing (that is, $\alpha = 1$ and $0 < p < 1$). The supervisor's "soft" signal is about the agent's high productivity.

7 Conclusion

This paper shows that in a principal-supervisor-agent relationship with collusion, the principal is better off with hard-but-non-forgeable information than with soft information when there are effort distortions and noisy signals. Only in the limit case of accurate signals the principal with soft information is as well off as with hard information. Effort distortions are needed for the creation of collusion stakes through differential information rents when the supervisor's signal is noisy.

The conditions under which the supervisor with soft information is still valuable for the principal are, first, that the supervisor's signal's must exceed some lower treshold of noise and, second, that side transfers between the agent and the supervisor must be inefficient. Finally, although this paper deals with supervision (i.e., *ex ante* signals) the exactly same results apply to auditing (see Cont [4]).

Appendix

Proof of Proposition 2 and Corollary 2: In a hard-and-non-forgeable-information environment, a feasible contract must satisfy the agent's participation and incentive constraints (2), the supervisor's participation/liability constraints (3) and the coalition constraints (8) and (9). We repeat the last two constraints here for convenience.

$$(H) \quad \begin{array}{c} Pr(\theta_H|\sigma = H)\,[w_{HH} + kR_{HH}] + Pr(\theta_L|\sigma = H)\,[w_{LH} + kR_{LH}] \geq \\ Pr(\theta_H|\sigma = H)\,[w_{H\emptyset} + kR_{H\emptyset}] + Pr(\theta_L|\sigma = H)\,[w_{L\emptyset} + kR_{L\emptyset}] \end{array} \tag{14}$$

$$(L) \quad \begin{array}{c} Pr(\theta_H|\sigma = L)\,[w_{HL} + kR_{HL}] + Pr(\theta_L|\sigma = L)\,[w_{LL} + kR_{LL}] \geq \\ Pr(\theta_H|\sigma = L)\,[w_{H\emptyset} + kR_{H\emptyset}] + Pr(\theta_L|\sigma = L)\,[w_{L\emptyset} + kR_{L\emptyset}] \end{array} \tag{15}$$

where the probabilites correspond to the supervisor's better information after having observed the signal (updated beliefs)

$$\begin{aligned} Pr(\theta_H|\sigma = H) &= \frac{q\alpha}{q\alpha + (1-q)(1-\alpha)} \qquad Pr(\theta_L|\sigma = H) = \frac{(1-q)(1-\alpha)}{q\alpha + (1-q)(1-\alpha)} \\ Pr(\theta_H|\sigma = L) &= \frac{q(1-\alpha)}{q(1-\alpha) + (1-q)\alpha} \qquad Pr(\theta_L|\sigma = L) = \frac{(1-q)\alpha}{q(1-\alpha) + (1-q)\alpha} \end{aligned}$$

We simplify these constraing in three steps. First, racall that the solution of the optimal contract udner no collusion the principal leaves no rent to the less productive agent. This result still holds when there is threat of collusion, i.e., $t_{Lr}^{NFI} = \left(e_{Lr}^{NFI}\right)^2/2$ for $r = \emptyset, L, H$. Second, constraints IC(Hr) are binding (standard result), and $t_{Hr}^{NFI} = \left(e_{Hr}^{NFI}\right)^2/2 + R_{Hr}^{NFI}$ with $R_{Hr}^{NFI} = R\left(e_{Lr}^{NFI}\right)^2/2$. Third, since the agent earns no rent in states L∅, LL and LH, the principal pays $w_{L\emptyset}^{NFI} = w_{LL}^{NFI} = w_{LH}^{NFI} = 0$ to the supervisor (the coalition does not benefit from changing the supervisor's report given that productivity is low). Using all these results in (14) and (15), we have that

$$w_{HH} + kR_{HH} \geq w_{H\emptyset} + kR_{H\emptyset} \tag{16}$$

$$w_{HL} + kR_{HL} \geq w_{H\emptyset} + kR_{H\emptyset} \tag{17}$$

that is, the coalition conditions simplify to ensuring that there are no group incentives to conceal the supervisor's information given that the agent is productive. Moreover, the simplified coalition constraints are *as if* the supervisor knew that the agent's type is θ_H when bargaining with the agent over a side-contract.[15]

Given the simplified coalition constraints, it is clear that when the supervisor observes nothing ($\sigma = \emptyset$) and the agent is productive the principal sets $w_{H\emptyset}^{NFI} = 0$ (the superivsor has no information and cannot forge it).[16]

Next, we solve the restricted optimization problem with constraint CC(HH) binding, and then check that the solution to this problem is the solution to the unrestricted problem because constraint (17) with $w_{HL} = 0$ is never binding.

Consider first $p < 1$. The principal's problem simplifies to maximize

$$\sum_{r \in \{0,L,H\}} \left\{ \pi_{Hr} \left[\theta_H e_{Hr} - \frac{e_{Hr}^2}{2} - R\frac{e_{Lr}^2}{2} \right] + \pi_{Lr} \left[\theta_L e_{Lr} - \frac{e_{Lr}^2}{2} \right] \right\} - \pi_{HH} Rk \left(\frac{e_{L\emptyset}^2}{2} - \frac{e_{LH}^2}{2} \right)$$

From the first-order conditions for a maximum we have

$$e_{H\emptyset}^{NFI} = e_{HL}^{NFI} = e_{HH}^{NFI} = \theta_H \qquad e_{L\emptyset}^{NFI} = \frac{(1-q)(1-p)\theta_L}{(1-q)(1-p) + qR\left[(1-p(1-\alpha k)\right]}$$

$$e_{LL}^{NFI} = \frac{(1-q)\alpha\theta_L}{(1-q)\alpha + qR(1-\alpha)} \qquad e_{LH}^{NFI} = \frac{(1-q)(1-\alpha)\theta_L}{(1-q)(1-\alpha) + qR\alpha(1-k)}$$

It is easy to check that $e_{LL}^{NFI} > e_{L\emptyset}^{NFI}$ and hence $R_{HL} > R_{H\emptyset}$. Constraint (17) holds strictly with $w_{HL} = 0$. On the other hand, $e_{L\emptyset}^{NFI} > e_{LH}^{NFI}$ (and then $R_{H\emptyset}^{NFI} > R_{HH}^{NFI}$) if and only if $\alpha > \alpha_h$, where

$$\alpha_h = \begin{cases} \dfrac{-[2(1-p)-k] + \sqrt{[2(1-p)-k]^2 + 4pk(1-p)}}{2pk} & \text{if} \quad k > 0 \\ 1/2 & \text{if} \quad k = 0 \end{cases} \tag{18}$$

When $\alpha \leq \alpha_h$, effort and compensations are such that $e_{LH}^{NFI} = e_{L\emptyset}^{NFI}$, $t_{HH}^{NFI} = t_{H\emptyset}^{NFI}$ and $w_{HH}^{NFI} = 0$ (that is, constraint(16) is binding with $w_{HH}^{NFI} = 0$), where

$$e_{L\emptyset}^{NFI} = e_{LH}^{NFI} = \frac{(1-q)(1-p\alpha)\theta_L}{(1-q)(1-p\alpha) + qR(1-p(1-\alpha))}$$

The principal's utility is

[15] In Tirole (1988) the supervisor knows the agent's type when he obtains a signal. This corresponds to $\alpha = 1$ in our model.

[16] As we discussed in the main body, there may exist incentives to blackmail the agent. We rule this possibility out. See footnote 9.

$$EU_P^{NFI} = \begin{cases} q\dfrac{\theta_H^2}{2} + \dfrac{(1-q)^2\theta_L^2}{2}\left\{ \begin{array}{c} \dfrac{p\alpha^2}{(1-q)\alpha + q(1-\alpha)R} + \\ \dfrac{(1-p\alpha)^2}{(1-q)(1-p\alpha) + q(1-p(1-\alpha))R} \end{array} \right\} & \text{if } \alpha \leq \alpha_h \\ \\ q\dfrac{\theta_H^2}{2} + \dfrac{(1-q)^2\theta_L^2}{2}\left\{ \begin{array}{c} \dfrac{p\alpha^2}{(1-q)\alpha + q(1-\alpha)R} + \\ \dfrac{p(1-\alpha)^2}{(1-q)(1-\alpha) + q\alpha(1-k)R} + \\ \dfrac{(1-p)^2}{(1-q)(1-p) + q\left((1-p(1-\alpha k)R\right)} \end{array} \right\} & \text{if } \alpha > \alpha_h \end{cases} \tag{19}$$

Consider now $p = 1$. Concealment is not possible and hence the coalition constraints are no longer relevant. The solution to the principal's problem is as in (5)-(6). The principal's utility is EU_P^{NC} from (7).

The maximizand is concave and second order conditions are satisfied.

From the results above, the principal always hires the supervisor. On the other hand, if k increases, ceteris paribus, α_h increases. If parameters are such that $\alpha < \alpha_h$, EU_P^{NFI} does not change with k. Otherwise, EU_P^{NFI} decreases in k. *Q.E.D.*

Proof of Proposition 3: The proof of the Proposition is similar to that of Proposition 2. First, we solve the optimal contract with the binding constraints IR(Lr), IC(Hr) and the coalition constraints (from equations (2), (3), (11) and (12)), and then we check the conditions under which the other constraints are strictly satisfied. Without loss of generality we can set $w_{Lr}^{FI} = 0$. Also, $w_{HL}^{FI} = 0$ since this is the most expensive state to pay the supervisor a collusion-proof wage. The principal's problem simplifies to maximize

$$\sum_{r\in\{0,L,H\}} \pi_{Lr}\left\{\theta_L e_{Lr} - \frac{e_{Lr}^2}{2}\right\} + \sum_{r\in\{0,L,H\}} \pi_{Hr}\left\{\theta_H e_{Hr} - \frac{e_{Hr}^2}{2} - R\frac{e_{Lr}^2}{2} - Rk\left(\frac{e_{LL}^2}{2} - \frac{e_{Lr}^2}{2}\right)\right\}$$

From the first-order conditions for a maximum we have (second-order conditins are satisfied):

$$e_{H\emptyset}^{FI} = e_{HL}^{FI} = e_{HH}^{FI} = \theta_H \qquad e_{LL}^{FI} = \frac{(1-q)p\alpha\theta_L}{(1-q)p\alpha + qR\left[k + p(1-\alpha)(1-k)\right]}$$

$$e_{L\emptyset}^{FI} = \frac{(1-q)\theta_L}{(1-q) + qR(1-k)} \quad e_{LH}^{FI} = \frac{(1-q)(1-\alpha)\theta_L}{(1-q)(1-\alpha) + q\alpha R(1-k)}$$

Note that $e^{FI}_{L\emptyset} > e^{FI}_{LH}$. Now, $e^{FI}_{LL} > e^{FI}_{L\emptyset}$ if and only if $\alpha > \alpha^f_2$, where

$$\alpha^f_2 = \frac{k + (1-k)p}{2p(1-k)} \tag{20}$$

When $\alpha = 1$ this contract corresponds to the contract with hard and non-forgeable information (Proposition 2).

The second case is $e^{FI}_{LL} = e^{FI}_{L\emptyset} > e^{FI}_{LH}$ (for $\alpha \leq \alpha^f_2$). Plugging this restriction into the program and solving for effort levels, we have e^{FI}_{LH} as before and

$$e^{FI}_{L\emptyset} = e^{FI}_{LL} = \frac{(1-q)(1-p(1-\alpha))\theta_L}{(1-q)(1-p(1-\alpha)) + qR\left[1 - p\alpha(1-k)\right]}$$

Next, $e^{FI}_{L\emptyset} = e^{FI}_{LL} > e^{FI}_{LH}$ if $\alpha^f_1 < \alpha \leq \alpha^f_2$, where

$$\alpha^f_1 = \frac{1}{2-k} \tag{21}$$

Otherwise, the solution is to set $e^{FI}_{L\emptyset} = e^{FI}_{LL} = e^{FI}_{LH}$, i.e., not to hire the supervisor, which corresponds to $\alpha \leq \alpha^f_1$. The payoff to the principal is

$$EU^{FI}_P = \begin{cases} EU^{NS}_P & \text{if } \alpha < \alpha^f_1 \\ q\dfrac{\theta^2_H}{2} + \dfrac{(1-q)^2\theta^2_L}{2}\left\{ \begin{array}{l} \dfrac{p(1-\alpha)^2}{(1-q)(1-\alpha) + q\alpha R\,(1-k)} + \\ \dfrac{(1-p(1-\alpha))^2}{(1-q)(1-p(1-\alpha)) + qR(1-p\alpha(1-k))} \end{array} \right\} & \text{if } \alpha^f_1 \leq \alpha < \alpha^f_2 \\ \dfrac{\theta^2_H}{2} + \dfrac{(1-q)^2\theta^2_L}{2}\left\{ \begin{array}{l} \dfrac{(1-p)}{(1-q) + qR\,(1-k)} + \dfrac{p(1-\alpha)^2}{(1-q)(1-\alpha) + q\alpha R\,(1-k)} \\ + \dfrac{p^2\alpha^2}{(1-q)p\alpha + qR\,(k + p(1-\alpha)(1-k))} \end{array} \right\} & \text{if } \alpha^f_2 \leq \alpha < 1 \\ EU^{NFI}_P & \text{if } \alpha = 1 \end{cases} \tag{22}$$

Q.E.D.

Proof of Theorem 1.[17]

a) Supervisor: The agent's information corresponds to her type and the supervisor's signal. The supervisor's information is his signal.

[17] The proof follows Baliga's [2] steps.

Consider the following direct mechanism: The principal asks a message report $a = (a_A, a_\sigma) \in \{\theta_L, \theta_H\} \times \{\emptyset, \theta_L, \theta_H\}$ to the agent and a message report $r \in \{\emptyset, \theta_L, \theta_H\}$ to the supervisor, and sets allocations and compensations according the following rule: $\rho(m) = \{e_m, t_m, w_m\}$ (where $m = (a, r)$, and e, t, and w are from Proposition 3,[18] such that

$$\rho(m,x) = \begin{cases} \{0,0,0\} & \text{if } a_\sigma \neq r \\ \{e^{FI}_{Lr}, t^{FI}_{Lr}, 0\} & \text{if } a = (\theta_L, r) \text{ and } r \in \{\emptyset, \theta_L, \theta_H\} \quad (a_\sigma = r) \\ \{e^{FI}_{Hr}, t^{FI}_{Hr}, k(t^{FI}_{HL} - t^{FI}_{Hr})\} & \text{if } a = (\theta_H, r) \text{ and } r \in \{\emptyset, \theta_L, \theta_H\} \quad (a_\sigma = r) \end{cases}$$

The agent exerts effort e such that $e \in argmax \quad t(m) - e^2/2$. Her strategy is $\zeta_A(\theta, \sigma) = (a(\theta, \sigma), e(\theta, \sigma))$. The supervisor's strategy is $\zeta_M(\sigma) = r(\sigma)$. Then $\zeta = (\zeta_A(\theta, \sigma), \zeta_M(\sigma))$.

Let a strategy profile ζ and set of beliefs ν be a Perfect Bayesian Equilibrium if players do not have incentives to change their strategy at any information set given beliefs, the other players' strategy, and beliefs are updated according to Bayes's rule whenever possible.

Let a side-transfer b from the agent to the supervisor be feasible for a given signal σ and manipulation of reports $m' = (a', r')$, if i) $t(m') \geq b \geq -w(m')$ and ii) the agent exerts effort $e' \in argmax \quad t(m') - e'^2/2$.

A strategy $s = (m')$ is individually profitable for the supervisor if $w(s) > w(\zeta|\sigma)$, and it is individually profitable for the agent if $t(s) - e'^2/2 > t(m(\theta, \sigma), \theta) - e(\theta, \sigma)^2/2$.

Let a collusive strategy $cs = (m', b)$ be *coalition profitable* for a signal σ and strategy ζ if a) it is feasible and b) the parties are *strictly better off*, i.e,

i) $t(m') - e'^2/2 - b > t(m(\theta, \sigma), \theta) - e(\theta, \sigma)^2/2$

ii) $w(m') + b > w(\zeta|\sigma)$

Definition: A strategy ζ is Collusion-Proof equilibrium if it is Perfect Bayesian Equilibrium and if there is no feasible and individually or coalition profitable strategy for supervisor and agent under any signal.

According to the allocation rule above, we show that the agent's and supervisor's truthful report of the private information and and the agent's acceptance of the principal's effort recommendation is an equilibrium strategy ζ that satisfies collusion proofness. On the one hand, neither the agent nor the supervisor has incentive to individually deviate. If their reports of the supervisor's signal differ ($a_\sigma \neq r$), they will get no utility. The

[18] The value the cost C is recovered from the type report and effort recommendation from that mechanism.

agent does not have incentives to change her type report (the rule satisfies participation and incentive constraints).

On the other hand, we show that the agent-supervisor coalition does not find a feasible deviation. When the agent's type is θ_L, she earns no rent and the supervisor's wage is 0. Any mutual change of their report of the supervisor's signal is not profitable. A possible deviation may involve the agent changing her report to θ_H and the coalition changing their report to any r. In this case the agent gets a negative rent since $t_{Hr} - e_{Hr}/2\Delta\theta < 0$ for every supervisor's message r. So this deviation could be possible if the supervisor pays the agent up to his wage.[19] The agent's utility in this case is

$$t_{Hr} - \frac{e_{Hr}^2}{2\Delta\theta} + kw_{Hr} = -\frac{R\theta_H^2}{2\Delta\theta} + (1-k^2)\frac{Re_{Lr}^2}{2} + k^2\frac{Re_{LL}^2}{2} < 0$$

for every supervisor's message r. Then any message change a' that involves the agent modifying her report from θ_L to θ_H is not profitable.

Suppose that the type-θ_H agent induces the supervisor to change their report of the supervisor's signal (to some $a'_\sigma = r'$). This is profitable from states H$\emptyset$ or HH to HL and from HH to H$\emptyset$. In all cases the agent gains the difference in her wage, but this is the exact amount needed to compensate the supervisor's wage reduction, which violates conditions i) and/or ii) of coalition profitability. *Q.E.D.*

References

[1] Antle, Rick, 1984, Auditor Independence, *Journal of Accounting Research*, vol. 22, No. 1, pp. 1-20.

[2] Baliga, Sandeep, 1999, Monitoring and Collusion with “Soft” Information, *Journal of Law, Economics and Organization*, vol. 15, No. 2, pp. 434-440.

[3] Baron, David and David Besanko, 1984, Regulation, Asymmetric Information and Auditing, *RAND Journal of Economics*, vol. 15, No. 4, pp. 447-470.

[4] Cont, Walter, 2001, Essays on Contract Design: 12345, UCLA Ph.D. Dissertation.

[5] Dalton, Melville, 1966, *Men Who Manage*, New York: John Wiley and Sons.

[19] We also assume that a side tranfer b from the supervisor is valued kb by the agent.

[6] Faure-Grimaud Antonie, Jean-Jacques Laffont and David Martimort, 1999, The Endogenous Transaction Costs of Delegated Auditing, *European Economic Review*, vol. 43, No. 4-6, pp. 1039-1048.

[7] Faure-Grimaud Antonie, Jean-Jacques Laffont and David Martimort, 2000, Collusion, Delegation and Supervision with Soft Information, *Review of Economic Studies*, vol. 70, pp. 253-279.

[8] Kofman, Fred and Jackes Lawarrée, 1993, Collusion in Hierarchical Agency, *Econometrica*, vol. 61, No. 3, pp. 629-656.

[9] Tirole, Jean, 1986, Hierarchies and Bureaucracies: On the Role of Collusion in Organizations, *Journal of Law, Economics and Organization*, vol. 2, No. 2, pp. 181-214.

[10] Tirole, Jean, 1992, Collusion and the Theory of Organizations. In J.J. Laffont (ed.), *Advances in Economic Theory*, vol. 2, Cambridge University Press, pp. 151-205.