MORAL HAZARD AND COLLUSION IN HIERARCHIES*

Charles Angelucci†    Antonio Russo‡

March 2012

Abstract

We investigate the impact of informal agreements between supervisors and supervisees on a firm’s optimal organization. In particular we study the extent to which administrative tasks should be delegated to supervisors. To do so we adopt the principal-supervisor-agent paradigm with a moral hazard problem at the bottom. Side-contracting between supervisor and agent takes place both before and after the latter has taken his action and information is soft. To deter collusion the supervisor must be offered a group-based compensation scheme. This removes his incentives both to engage in individual opportunism and collusive agreements with the agent. Cooperation, which takes the form of the supervisor treating supervisees fairly, occurs in equilibrium. Consequently, the principal may be better-off having the supervisor be a monitor rather than carrying only ex-post audits. Contrarily to the verifiable (hard) information case, the firm does

---

*We thank Philippe Aghion, Andrea Attar, Julia Cagé, Jacques Crémer, Emmanuel Farhi, Renato Gomes, Olga Gorelkina, Martijn Han, Oliver Hart, Jacques Lawarrée, Yassine Lefouili, Patrick Legros, Ching-to Albert Ma, Inès Maché-Stadler, David Martimort, Simone Menaglia, Humberto Moreira, Patrick Rey, François Salanié, Wilfried Sand-Zantman, Kouroche Vafaï and seminar participants in Alicante (ASSET 2010), Crest-Lei, Ecole Polytechnique, Paris-10, ECARES, Cargese (ESNIE 2011) and Oslo (EEA 2011) for helpful discussions and comments. Opinions and errors are ours.

†Toulouse School of Economics. Email: charles.angelucci@tse-fr.eu

‡Toulouse School of Economics. Email: antonio.russo@tse-fr.eu
not delegate payroll authority to the supervisor, although it does request reports concerning employee performance. Finally, we show that internal whistleblowing programs do not reduce the scope for harmful informal agreements.

JEL Classification: D82, D86, L23

Keywords: collusion, extortion, delegation, mechanism design
1 Introduction

Informal agreements - agreements that are not enforced by courts but sustainable because of trust, repeated relationships, violence, etc... - play a prominent role in the economy. In particular, and as emphasized by sociologists, hidden contracts between members of an organization are ubiquitous (see Roethlisberger and Dickson [1947], Dalton [1959] and Crozier [1963]). These aim at either cooperative or opportunistic ends and may differ in their broader impact. In this paper we investigate the optimal organizational response to informal agreements between supervisors and supervisees.

Cooperation between individuals is beneficial to the organization and may take many forms. In the model we develop a supervisor is said to be cooperative if she treats fairly her subordinate, despite other attitudes being more profitable (e.g. being tough). Opportunism, which may either be collective or individual, is instead detrimental. Consider collective opportunism: coalitions of individuals may emerge so as to enforce their own objectives. Such collusion may range from the simple exchange of favors to outright fraud, involving the hidden transfer of money or goods. Consulting firm KPMG for instance reports that over fifty per cent of corporate fraud in North America is made possible because of collusion between employees.\footnote{A series of surveys carried out by KPMG provide staggering figures: the average loss (to the firm) per fraud in North America is of $1.2 million. Over sixty per cent of corporate fraud involves employees and half of the time management is involved.} In our model the supervisor is tempted to cover up her subordinate’s shirking in exchange for a bribe or favor. Individual opportunism operates differently: it benefits only the instigator. A company’s supervisor - and the supervisor we consider in our model - may be tempted to report falsely the performance of subordinates if lucrative to do so even without side transfers of money (if for instance being tough is rewarded). In the extreme, a supervisor may even be tempted to engage in extortion: threaten to take an action affecting negatively her subordinate so as to extract a bounty. In this vein, Vafaï [2002, 2010] takes sexual harassment at the workplace as an example of such behavior.
Economics, since the seminal paper by Tirole [1986], has addressed some of these issues. In these models, as in ours, agents may reach binding, though hidden, agreements. One may invoke honor, friendship, and repeated relationships to justify their enforceability.\footnote{2} A large part of the existing literature studies adverse selection setups (see for instance Tirole [1986, 1992] or more recently Celik [2009]). We instead study a moral hazard framework. A principal hires an agent to work on a project and the latter unobservably chooses his effort level. While the principal wishes the agent to work hard, the latter prefers shirking. Although some important work has been carried out also in these environments (see for instance Holmstrom and Milgrom [1990], Itoh [1993] and Baliga and Sjostrom [1998]), the tendency has been to focus on agreements between agents performing similar tasks. Our focus is rather on vertical agreements.\footnote{3}

A supervisor is also hired, and her task is to report unverifiable information concerning the agent’s performance so as to reduce the latter’s information rent. The principal first offers a grand contract specifying payments to both the agent and the supervisor. Subsequently, the supervisor offers her own, hidden, side contract to the agent. No restrictions are imposed on the intent of side contracting: It may lead to either cooperation, collusion, or extortion. As a leading example of opportunism we consider the case of payroll fraud which, as the following citation suggests, fits well with the model:

“\text{(one) way to obtain approval of a fraudulent time card is to collude with a supervisor who authorizes timekeeping information. In these schemes, the supervisor knowingly signs false time cards and usually takes a portion of the fraudulent wage. In some cases, the supervisor may take the entire amount of the overpayment. In an example, a supervisor assigned employees to better work areas or better jobs, but in return demanded payment. (…)\text{}}
The employees were compensated for fictitious overtime, which was kicked back to the supervisor.”

Joseph T. Wells, chairman of the Association of Certified Fraud Examiners.

While the issue of collusion has been well investigated, that of extortion has been somewhat overlooked, with the exceptions of Khalil, Lawarrée and Yun [2010], hereafter KLY, and Vafaï [2010]. As KLY and Vafaï argue, both forms of opportunism impact differently the principal. While collusion reduces the agent’s gains from shirking (since a bribe needs to be paid not to be caught) extortion instead punishes the hard working agent (since a bribe needs to be paid not to be unfairly punished). Our model differs from theirs in that side contracting may take place both before (ex ante) and after (ex post) the agent has chosen his action (as opposed to only ex post). We believe this to be a particularly reasonable description of relationships within firms: these tend to foster close ties. Under ex ante side-contracting the supervisor may influence directly the agent’s action and, as a result, opportunism is potentially much more of a problem. We show that both rounds of side negotiations may differ in their impacts on the organization, and in particular on the usefulness of the supervisor’s information, and let the principal choose when to appoint the former. If the supervisor is present from the outset - both ex-ante and ex-post forms of opportunism are an issue - we say that she acts as a monitor. Her task is to observe and report on the agent’s behavior during the entire course of the latter’s activity. If instead she is appointed after the agent has chosen his effort level (and completed his task) - only ex post opportunism is an issue - she is then said to act as an auditor. Moreover, we further differentiate ourselves from KLY [2010] and Vafaï [2002, 2010] in that our supervisor’s reports are entirely

\[4\]

Other types of corruption, such as undisciplined drivers trying to bribe policemen, are probably better explained by ex post side contracting. However, our setup can also apply to prolonged monitoring of firms by public officials. An example can be the monitoring of environmental impacts of infrastructure projects by public agencies.
unverifiable (i.e. “soft”), thereby giving even further leverage to opportunism. To control for the fact that side-contracting may be costly to sustain in practice, we assume that hidden transfers possibly involve an exogenous transaction cost. Further on, while we initially let the principal communicate only with the supervisor, and exclusively concerning the realization of her private signal, we eventually relax all restrictions concerning communication. Finally, throughout the analysis, we investigate whether the principal can be better off by delegating the task of contracting with the agent, i.e. payroll authority, to the supervisor.

We now provide a summary of our main results.

**Cost of opportunism and auditing vs. monitoring.** The enforceability of the side contract, combined with the full manipulability of information, makes opportunism *a priori* a very serious threat to the well functioning of the firm. Under similar assumptions, but in adverse selection models, some papers nevertheless establish that side contracting may be harmless (see Laffont and Martimort [1997, 1998] and Che [2006]). In contrast to them we find that informal agreements are always costly. This is because they reduce the principal’s ability to enjoy the supervisor’s superior information concerning the performance of the agent. In a nutshell, under adverse selection, as opposed to moral hazard, the principal is able to better exploit the inefficiencies stemming from information asymmetries within the coalition.

In particular, if the supervisor is appointed *ex ante* (i.e. as a monitor) we find that collusion, as opposed to extortion, constitutes the most severe threat to the functioning of the organization. This result stands in stark contrast to those in KLY [2010] and Vafaï [2002, 2010]. Essentially, because in these models side contracting occurs *ex post,*

---

*Soft information certainly looks an important environment to consider. Examples of supervisory reports being unverifiable abound. As an example, in his study of the “Milo Fractionating Plant”, Dalton (1959), reports cases of costs being exaggerated by creation of fictitious personnel or overstatement of costs of equipment. In KLY [2010] information is soft only under collusion. Faure-Grimaud Laffont Martimort [2000, 2003a, 2003b] also consider soft information. Baliga (1999) considers Tirole (1992)’s three-tier hierarchy with soft instead of hard information and shows that agency costs are unchanged.*
the supervisor cannot be made to internalize the consequences of opportunism. In our model, however, to prevent collusion, the optimal incentive scheme is such that it rewards both players when the project is successful. It then becomes in the supervisor’s best interest to act cooperatively and induce the agent into exerting high effort. This, we show, necessarily involves a promise not to engage in extortion ex post. Preventing collusion implies preventing extortion. When the supervisor is appointed ex post (i.e. as an auditor) we then find, in line with KLY [2010] and Vafaï [2002, 2010], that extortion is the most harmful form of opportunism. This amounts to saying that the principal must choose between collusion or extortion when deciding whether to appoint an auditor or a monitor. We find that if transaction costs of side transfers are sufficiently high it is then optimal to appoint the supervisor early. Indeed, because of the threat of extortion, it is optimal to never make full use of the information she collects. This comes at the cost of blunting incentives for the agent. Appointing the supervisor early can make the organization avoid this problem. However, extra rewards must then be paid to her, to prevent collusion. These rewards are decreasing, ceteris paribus, in the transaction costs of side-transfers: as long as these are sufficiently high, we find, the gains of appointing the supervisor early outweigh the costs. Otherwise it is optimal to appoint the supervisor at a later stage.

The decision to appoint the supervisor at an early or at a later stage can be recast in terms of designing an organization in which the supervisor functions as a monitor rather than one in which he functions as an auditor. Finding that an organization could be better off by fostering ties between supervisors and supervisees runs somewhat against common wisdom. It is instead usually thought that auditing tasks should be delegated to frequently changed auditors. Intuitively these arrangements aim at diluting the capacity to develop privileged relationships, thought to ease collusion. Our result

---

6 Strausz [2006] considers a similar question. However, he studied a model in which information is gathered directly by the principal: there is no supervisor. We can however borrow his terminology: “The difference between the two procedures is that monitoring takes place while the agent chooses his action whereas auditing occurs after he has taken his action” (Strausz, 2006, p.1).
suggests a drawback of such schemes: the same factors facilitating collusion may also foster cooperation. By appointing a supervisor for a long period of time the organization is able to make her an essential input in the project. Only by doing so can she be made to internalize the consequences of opportunism on the performance of the firm.

**Augmented mechanisms.** The results discussed in the previous paragraph were obtained under the assumption that the principal can communicate only with the supervisor, and exclusively regarding her private signal concerning the agent’s choice of effort. While reasonable for many applications in mind, this might be too restrictive and give rise to some concerns on theoretical grounds. To see this note that both supervisor and agent have additional private information of interest to the principal, say the nature of the side contract, and thus one cannot rule out a priori more complex mechanisms that would attempt at eliciting it. In the language of the common agency literature, supervisor and agent are said to have market information. We thus eventually relax all restrictions on what the principal can do and find that it can indeed be better off by communicating more extensively with its subordinates. In particular we show that the optimal grand contract involves the design of a self reporting mechanism that elicits the status of ex ante side negotiations. This is an important tool at the disposal of the principal by which to manipulate the bargaining strength of the agent vis-à-vis the supervisor, and therefore the profitability of collusion to the latter. Under such a scheme, we find an equivalence between ex ante and ex post side contracting: The principal is indifferent as to appoint an auditor or a monitor, even in the absence of transaction costs of side transfers. In other words, we show that the principal is never worse off by appointing a supervisor early if appropriate communication schemes are designed.

**Decentralized contracting.** Throughout the analysis, and in both versions of the model, we investigate whether the optimal organizational response to opportunism
involves delegating payoff authority to the supervisor, i.e. essentially creating an arms' length relationship. We find that because of a double marginalization of rents an arms' length relationship is always strictly suboptimal. Note that this is true even if the communication channel with the agent remains open. This result is in contrast to adverse selection models in which there is usually no loss of generality in relying on the supervisor to contract exclusively with the agent (See for instance Faure-Grimaud et al. [2003b]). In moral hazard set ups with lateral side-contracting Holmstrom and Milgrom [1990] and Itoh [1993] show that letting side-contracting occur is strictly optimal. In a similar set-up, but under risk neutrality, Baliga and Sjostrom [1998] show that the optimal centralized contract can be implemented through decentralization.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 solves the game first allowing supervisor and agent to side-contract only once the latter has chosen his action, and then introducing the possibility of side-contracting also before that. It ends by deriving implications for delegation of payroll authority to the supervisor and for the optimal organization of internal audits. Section 4 studies the role of internal whistleblowing programs. Section 5 concludes.

2 The model

2.1 Information and Players

A principal $P$ (the organization), contracts both with a supervisor $S$ (she), and an agent $A$ (he). All players are assumed risk neutral. $A$ and $S$'s have zero reservation utilities. None of them has private wealth and both are protected by limited liability.

\footnote{In an adverse selection environment, Celik [2009] provides conditions under which full delegation, in the sense that the principal stops communicating with the agent, is suboptimal.}

\footnote{Macho-Stadler and Perez-Castrillo [1998] investigate issues of delegation when the principal suffers from commitment issues and provide conditions under which a decentralized structure is equivalent to a centralized one with collusion. Also, Che [1995] and Alger and Ma [2003] study models in which tolerating collusion is best.}
Information. There is a publicly available signal $\pi$ and a signal $\sigma$ observed only by $S$ and $A$. Both are correlated with $A$’s action in a way that will be shown below. Finally, we have the publicly observable message $m$ that $S$ makes to $P$ concerning $\sigma$. Note that in this first version of the model we assume communication to take place only with $S$, and exclusively concerning the realization of signal $\sigma$. All restrictions on communication are relaxed in section 4. We assume that $\sigma$ is soft information, in the sense that $m$ can take any value on the support of $\sigma$, regardless of its actual realization.

We define a state as the combined realization of $\pi, \sigma$ and $m$.

Agent. A unobservably chooses a binary action $e \in \{e, \bar{e}\}$. This action may be given various interpretations, but, to fix ideas, we interpret it as the choice between two levels of effort on the job: low $e$ and high $\bar{e}$. Effort $e$ implies disutility $\psi(e)$ where $\psi(\bar{e}) = \psi > \psi(e) = 0$. Signal $\pi$ may take one of two values: high $\pi = \pi$ or low $\pi = \bar{\pi}$ and is correlated with $e$. It may be interpreted, for instance, as the profitability of the project $A$ works on. Its conditional distribution is as follows:

<table>
<thead>
<tr>
<th></th>
<th>$e = \bar{e}$</th>
<th>$e = e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\pi}$</td>
<td>$\rho_\pi$</td>
<td>$1 - \rho_\pi$</td>
</tr>
<tr>
<td>$\bar{\pi}$</td>
<td>$1 - \rho_\pi$</td>
<td>$\rho_\pi$</td>
</tr>
</tbody>
</table>

where $\rho_\pi > \frac{1}{2}$; i.e. $\bar{\pi}$ ($\pi$) is more likely when high (low) effort is made. $A$ receives a state-contingent transfer $t_{\pi m} \geq 0$ from $P$ and makes/receives side-transfer $y_{\pi \sigma m}$ from/to $S$. The latter is assumed to be positive when going from $S$ to $A$ and negative otherwise. $A$’s utility function is $U^A = t_{\pi m} + y_{\pi \sigma m} - \psi(e)$.

Supervisor. With probability $p$, $S$ privately observes a signal $\sigma$ that is perfectly informative of $e$. The signal is empty with probability $1 - p$. More precisely, the conditional distribution of $\sigma$ is:

\[\text{Transfers from } P \text{ can clearly be conditioned only on what it actually observes.}\]
We will refer to $G$ as “good news” on $A$’s effort, $B$ as “bad news” and $N$ as “no news”. $S$’s utility function is $U^S = s_{\pi m} - K(y_{\pi \sigma m})y_{\pi \sigma m}$, where $s_{\pi m} \geq 0$ and $y_{\pi \sigma m}$ denote respectively the state-contingent salary received from $P$ and the side-transfer made to/received from $A$. $K(y_{\pi \sigma m})$ captures the transaction cost of organizing side-transfers for $S$. We have $K(y_{\pi \sigma m}) = k$ if $y_{\pi \sigma m} < 0$ and $K(y_{\pi \sigma m}) = \frac{1}{k}$ otherwise. If $S$ wants to send 1 dollar to $A$, she has to pay $k > 1$ dollars, while if $A$ sends her 1 dollar, $S$ receives only $\frac{1}{k} < 1$ dollars.\(^{10}\)

**Principal.** Total (state contingent) payments by $P$ are denoted $z_{\pi m} = t_{\pi m} + s_{\pi m}$. Its utility function is $U^P = u(\pi) - z_{\pi m}$. We assume that $P$ gets some utility $u(.) > 0$ when $\pi = \bar{\pi}$ and none otherwise. Such utility is high enough that the organization always prefers to induce $A$ into choosing $\bar{e}$. This is why $P$’s objective is to induce high effort at the lowest possible cost. Such cost is measured by $E(z) = \sum_\pi \sum_\sigma p^e_{\pi \sigma} (t_{\pi m} + s_{\pi m})$, where $p^e_{\pi \sigma}$ is the probability, conditional on $e$, of state $(\pi, \sigma, m)$ taking place.\(^{11}\) It is useful to explicitly write the joint distribution of probabilities $p^e_{\pi \sigma}$

\(^{10}\) $S$ is not part of the production process. This differentiates our setting from other papers also focusing on moral hazard environments, in which $S$ has a role in production. As emphasized by Radner (1992) it is very commonly observed that management is a very distinct activity from production.

\(^{11}\) Since $m$ is not stochastic, $p^e_{\pi \sigma m}$ is invariant with respect to it. This is why we drop the index.
we will also denote the expected transfers to $A$ and $S$, as well as side transfers, conditional on the effort $e$, on the actual realization of $\sigma$ and on $m$, respectively as

$$E_{\sigma t m}^e \equiv \sum_\pi p_{\pi\sigma}^e t_{\pi m} \quad E_{\sigma s m}^e \equiv \sum_\pi p_{\pi\sigma}^e s_{\pi m} \quad E_{\sigma y m(\sigma)}^e \equiv \sum_\pi p_{\pi\sigma}^e y_{\pi\sigma m} \quad \forall e, \sigma, m$$

2.2 Contracts

Grand-Contract. $P$ designs a grand contract $GC$, which is a take-it-or-leave-it offer specifying payments $t_{\pi m}$ to $A$ and $s_{\pi m}$ to $S$. These can be contingent only on the information available to $P$.

Side-Contracts. We assume $S$ has full bargaining power in designing side contracts with respect to $A$ (due to her superior position in the hierarchy). A side-contract is a take-it-or-leave-it offer specifying side-transfers $y_{\pi\sigma m}$ and a reporting function $m = m(\sigma)$, conditional on $\sigma$. Side transfers are contingent on the realization of profits $\pi$, private signal $\sigma$ and message $m$. The commitment on $m(\sigma)$ is credible if and only if the report is interim rational given side-transfers.\footnote{S can commit only on side-transfers, which, as specified above, are contingent both on $r$ and $\sigma$. It is therefore by committing to a given set of side-transfers that $S$ determines the interim rationality of sending a given report for a given $\sigma$. This assumption is fully in line with existing literature. If $S$ could directly commit to any $m(\sigma)$ results would be almost identical.} In each state, money can flow from or to $A$, but him and $S$ may not exchange more than what is specified in the $GC$, i.e. $-t_{\pi m} \leq y_{\pi\sigma m} \leq s_{\pi m}$ for $\forall \pi, \sigma, m$.

2.3 Timing

The sequence of events is as follows

1. $P$ offers a grand contract $GC$ to $S$ and $A$

2. $S$ offers side-contract $EASC$ to $A$. We refer to this as the “ex ante” side-contracting stage
3. A chooses $e$. Signal $\sigma$ is realized

4. $S$ offers side-contract $EPSC$ to $A$. We refer to this as the “ex post” side-contracting stage

5. Report $m$ is made

6. Profits $\pi$ are realized

7. $GC$, $EASC$ and $EPSC$ are executed

We allow for two distinct side-contracting stages. The first takes place before $A$ has chosen his action (Stage 2). We refer to the result of such stage as an Ex Ante Side-Contract ($EASC$). The second takes place after $A$ has chosen his action and $\sigma$ is realized (Stage 4). We refer to the result of such stage as an Ex Post Side-Contract ($EPSC$).

Once the $EASC$ has been signed, the $EPSC$ may replace it for the relevant contingencies (i.e. for the value of $\sigma$ that has realized after $A$ has chosen $e$). In other words, we allow for renegotiation of the $EASC$. However, we assume collective renegotiation. For the given realization of $\sigma$, both $A$ and $S$ have to receive a weakly higher expected payoff under the new ($EPSC$) than under the initial side-contract ($EASC$). Both $A$ and $S$ can also decide to not participate to any of the two side-contracting rounds.

3 Optimal contracts

We here study the optimal organization under different assumptions on the timing of side-contracting. First, we provide two benchmark $GC$s that will be useful references for the subsequent analysis. Then, we consider the optimal $GC$ if side contracting is allowed only at Stage 4. Finally, we allow for side contracting both at Stage 2 and Stage 4.

$^{13}$If $S$ had the possibility of reneging unilaterally on the $EASC$, side-contracting would simply collapse to the ex post outcome, i.e. Stage 2 would always be void.
3.1 Benchmarks

No supervision. Suppose $P$ does not hire $S$. In that case, $\sigma = N$ always. It is then straightforward to show that the best $P$ can do is to offer a contract specifying a positive transfer to $A$, equal to $\frac{\psi}{(2\rho_s-1)}$, if and only if $\pi = \bar{\pi}$. Total expected payments are then $E(z)^{SB} = \frac{\rho_s \psi}{(2\rho_s-1)}$. We refer to this benchmark as the second-best (SB) contract. It is intuitive that in no circumstances will $P$ hire $S$ and design a GC such that $E(z) \geq E(z)^{SB}$.

Benevolent Supervision. Suppose now that $S$ is benevolent towards $P$, i.e. always reports information truthfully, so that $m = \sigma$. The incentive compatibility constraint at the agent’s level then takes the form:

$$\sum_{\sigma} E^e_\sigma t_\sigma - \psi \geq \sum_{\sigma} E^e_\sigma s_\sigma. \quad (1)$$

In the optimal $GC$, $P$ optimally sets $t_{BG} = \frac{\psi}{\rho_s}$ and all other transfers to $A$ and $S$ at zero. Since a perfectly informative signal on $e$ is available at no cost, the $GC$ allows $P$ to reach the first-best level of expected payments: $E(z)^{FB} = \psi$. We refer to this benchmark as the benevolent supervisor case.

3.2 Opportunistic supervisor with only ex-post side-contracting

Suppose the supervisor is no longer benevolent but opportunistic: i.e. she can engage in side-contracting with $A$ and make the report $m$ that maximizes her payoff. Assume, for the moment, that side-contracting occurs only ex-post (i.e., Stage 2 is void). Suppose to be at Stage 4. We proceed under the following

**ASSUMPTION:** In the absence of side agreements with $A$, if $\sigma = x$ and $E^e_\sigma s_x = \max(E^e_\sigma s_{x-})$, $S$ chooses $m(x) = x$.

This behavioral assumption, standard in the mechanism design literature, can cap-
ture an (infinitesimal) moral cost of fabricating evidence by $S$. It should be intuitive that $S$ has strong discretionary power given the nature of her private information and her bargaining power with respect to $A$. There are two ways in which she can exploit her position. The first is to try to extract money (or favors) from $A$. $S$ may propose $A$ a report $m(\sigma) = m$ while threatening, if he does not agree, to make an unfavourable report $m(\sigma) = l$. This allows her to ask for $y_{\pi m(\sigma)} = t_x t - t_{ym}$, extracting $A$’s entire willingness-to-pay to have $m$ instead of $l$ being reported. However, in order for this to work, reporting $l$ must be a credible threat, i.e. it has to be interim rational for $S$ to do so in the absence of an agreement with $A$. This is true only if $m(\sigma) = l$ strictly maximizes the expected salary $E^w_\sigma s_m(\sigma)$; or, if it weakly maximizes it, only if $l = \sigma$ (by the infinitesimal cost of fabricating evidence). Otherwise, $l$ is simply not a credible menace.

To make an example, $S$ can choose $m(\sigma) = l = B$ only if $E^w_\sigma s_B > \max (E^w_\sigma s_G, E^w_\sigma s_N)$. If $E^w_\sigma s_B = \max (E^w_\sigma s_G, E^w_\sigma s_N)$, then $l = B$ is a credible threat only if $\sigma = B$.

The second way to exploit her position is for $S$ to seek the highest salary paid by $P$, choosing $m(\sigma)$ to maximize $E^w_\sigma s_m(\sigma)$, without any side-transfer (so $y_{\pi m(\sigma)} = 0$). Note that both of the mentioned options are available whatever the true realization of $\sigma$, $S$’s information being fully soft. To conclude, $S$ will choose $m$, $l$ (where applicable) and the (possibly null) side-transfers $y_{\pi m(\sigma)}$ to maximize $E^w_\sigma \left(s_r - \frac{1}{k}y_{m(\sigma)}\right)$.

Quite clearly, given $S$’s bargaining power, $A$ can never benefit from side-contracting with her. At best, he is left with the same payoff as if $\sigma$ were truthfully reported, in cases where $l = \sigma$ (with $S$ capturing all the eventual extra transfers paid by $P$).

Let us now look at how $P$ can respond to such opportunistic behavior by $S$. The optimal GC is the solution to

\[ 14 \text{In case } E^w_\sigma s_G > \max (E^w_\sigma s_B, E^w_\sigma s_N) A \text{ could benefit from } S \text{ always reporting } G \text{ independently. However, such an ordering of report-contingent expected salaries to } S \text{ can never be optimal for } P, \text{ as it would make her information completely useless.} \]
\[
\min_{\{t_{\pi m}, s_{\pi m}\}} \sum_{\sigma} E_{\sigma}^e (s_{m(\sigma)} + t_{m(\sigma)}) \quad \text{s.t.} \quad \sum_{\sigma} E_{\sigma}^e (t_{m(\sigma)} + y_{m(\sigma)}) - \psi \geq \sum_{\sigma} E_{\sigma}^e (t_{m(\sigma)} + y_{m(\sigma)}).
\]

Constraint \( (2) \) being \( A \)'s incentive compatibility constraint, taking into account side-transfers exchanged with \( S \) and conditional on her report \( m(\sigma) \).

To see how this problem is solved, consider, first of all, that \( P \) must necessarily have \( E_{\sigma}^e s_G = E_{\sigma}^e s_B = E_{\sigma}^e s_N \). The reason is that, if any of the values among \( \{G, N, B\} \) guaranteed \( S \) a higher salary than the others, \( A \) would be left with the same expected transfer, whatever the realization of \( \sigma \). Either \( S \) makes always the report maximising \( E_{\sigma}^e s_m \), or she threatens \( A \) to do it. In either case, \( A \) is left with the expected transfer \( P \) would make if that report were made (a formal proof of this is provided in the Appendix). This would make \( S \)'s signal useless in order to provide incentives to \( A \). However, setting \( E_{\sigma}^e s_G = E_{\sigma}^e s_B = E_{\sigma}^e s_N \) still implies that the use made of \( S \)'s information is limited. In fact, it is easy to see that, as long as \( E_{\sigma}^e t_G > \max (E_{\sigma}^e t_B, E_{\sigma}^e t_N) \) she will always report \( G \), capturing bribes from \( A \) when possible (i.e. when \( \sigma \neq G \), given the above assumption). Thus, we have

\[
m(\sigma) = G \quad \forall \sigma \quad E_{\sigma}^e y_G = E_{\sigma}^e (t_{\sigma} - t_G),
\]

so that the expected net transfer \( E_{\sigma}^e (t_{m(\sigma)} + y_{m(\sigma)}) \) received (after paying the bribe to \( S \)) by \( A \) is \( E_{\sigma}^e t_{\sigma} \). \( P \)'s problem writes then as

\[
\min_{\{t_{\pi m}, s_{\pi m}\}} \sum_{\sigma} E_{\sigma}^e (s_G + t_G) \quad \text{s.t.} \quad \sum_{\sigma} E_{\sigma}^e t_{\sigma} - \psi \geq \sum_{\sigma} E_{\sigma}^e t_{\sigma}
\]

\[\text{15} A \text{ is protected by limited liability with respect to both } P \text{ and } S. \text{ This and the assumption of a zero outside utility for } A \text{ imply that } (2) \text{ also ensures participation to the } GC.\]
It is intuitive that, because $G$ will always be reported on the equilibrium path, $P$ is better off having $E^e_G t_N = E^e_G t_G$ in order to minimize expected payments. This implies that $G$ and $N$ are effectively treated as the same piece of information. On the other hand, since $B$ never turns up in equilibrium, $t_{\pi B} = 0$ is optimal. To continue, $P$ will optimally make payments only when $\pi = \bar{\pi}$, making use of the public signal. Finally, in order to minimize payments to $S$, $P$ can simply set $E^e_G s_G = E^e_G s_B = E^e_G s_N = 0$. The optimal GC is summarized in the following

**Proposition 1** Suppose $S$ and $A$ can side-contract only after the latter has chosen the level of effort $e$ and the supervisor has obtained signal $\sigma$ concerning such effort (i.e. Stage 2 is void). The optimal GC is

$$
I : t_{\pi G} = t_{\pi N} = \frac{\psi}{2\rho_\pi - 1 + (1 - \rho_\pi) p} \quad t_{\pi B} = 0 \forall \pi \quad t_{2m} = 0 \forall r \quad s_{\pi m} = 0 \forall \pi, m
$$

The total expected payment by $P$ is equal to $E(z)^I = \frac{\rho_\pi \psi}{2\rho_\pi - 1 + (1 - \rho_\pi) p}.$

**Proof.** See Appendix

**Discussion of the result.** Due to the supervisor’s strong discretionary power, the principal is forced to avoid conditioning her salary on the information she reports. Therefore, the firm has to give up on eliciting truthful reports. Bearing in mind the optimal organization in the “benevolent supervisor” case, it is clear that there is a cost in doing so. This is because it becomes impossible to discriminate between $G$ and $N$. Yet, not differentiating payments to $S$ according to $m$ allows $P$ to protect $A$ against her opportunistic behavior. Moreover, the cost of giving up on truthful reporting is limited. On the one hand, signal $B$ does not turn up in equilibrium. On the other, it is still as unattractive to $A$ as it would be in the case of benevolent supervision since, even if $G$ is reported instead, $A$ is still punished through the bribes captured by $S$.

---

16Grand Contract $I$ may recall the “Least-Cost-Corruption-Proof” contract of KLY (2010, Lemma 1). For them, anyway, such contract is not the optimal one. They obtain that, conditionally on $\pi$ being...
The inefficiency generated by $S$’s discretionary power, compared to the benevolent supervisor benchmark, is larger the less precise $S$’s information (i.e., the smaller is $p$). This is due to the fact that $P$ has to give up on distinguishing between states in which a signal that may turn up only with $A$ making high effort ($G$) and one which may also turn up when $A$ shirks ($N$) are reported. The more frequently $N$ turns up, the higher the cost in doing so.

KLY (2010) and Vafai (2010) point out that collusion (i.e. the supervisor using her discretionary power to make reports that benefit the agent) has a cost for the agent that shirked, while extortion (i.e. the supervisor using her power in a way that harms the agent) punishes the agent that took the right action, thereby weakening incentives. The results obtained here are in line with this view. Indeed, when only ex-post side-contracting is allowed, the threat of extortion implies that the organization cannot make full use of $S$’s information.

To conclude this section, it is important to note that $GC I$ is such that, in equilibrium, $S$ is left with a smaller rent than the one she might get should $A$ shirk. In fact, when $P$ offers $GC I$, $S$ actually gets no rent at all, since shirk takes place only out of equilibrium. If $A$ shirked, instead, she could capture the bribe $A$ would pay her when $B$ turns up. Yet, $S$ cannot break $A$’s indifference between high and low effort, since side-contracting happens only at the ex-post stage. At that point, $A$ has already chosen his action. We will show in the following that if $A$ and $S$ can side-contract before the former has chosen his action, $P$ then needs to make sure that $S$ does not induce the indicative of low effort by $A$ (i.e. $\pi = \bar{\pi}$), it is optimal to tolerate collusion when $\sigma = N$, while setting $t_{\pi G} > t_{\pi N} = 0$. Since $t_{\pi N}$ turns up more often when the agent shirks than when he does not, collusion punishes shirk. When $\pi = \bar{\pi}$ and $t_{\pi G} > t_{\pi N}$, however, the same type of collusion would punish the agent that works hard. Being conditional on $\pi = \bar{\pi}$, $t_{\pi N}$ is more likely to be paid when $e = \bar{e}$ than when $e = e$. Therefore, raising it all the way to $t_{\pi G}$ improves incentives for the agent. Also, in our setup, $t_{\pi G} = t_{\pi N} = 0$ is optimal, so there can be no scope for collusion when $\sigma = N$.

However, the authors concentrate on extortion concerning reports of good performance, since they consider “hard” supervisory information (at least if supervisor and agent do not collude). In contrast, we look at information that is soft, so the supervisor can credibly report proofs of shirk even if they did not turn up.
wrong one so as to increase her payoff in such a manner. This may radically change
the optimal incentive scheme.

3.3 Opportunistic supervisor: ex ante and ex post side contracting

We now study the situation in which $S$ can side-contract with $A$ both before and after
the latter has decided his effort level $e$. That is, side contracting happens at both Stage
2 and 4. As argued in Section 2, the side-contract signed at the ex-ante stage can be
renegotiated ex-post, for the relevant contingency (i.e. for the given value of $\sigma$ that has
turned up). However, we proceed assuming that this may happen only under collective
renegotiation. Under this condition, we can claim the following.

CLAIM: If the EASC can be renegotiated only under collective renegotiation,
then $S$ will offer an EASC that is renegotiation proof at the ex-post stage.

The claim is proved in the Appendix. What this implies is that there is no loss of
generality in considering EASC and EPSC as substitutes. If an EASC is signed, then,
we can disregard Stage 4.

$S$ has the option not to offer any EASC at all. However, since at Stage 2 she can
design all the side-contracts that are feasible at Stage 4, there is no gain for her in
unilaterally doing so. Yet, $A$ can refuse the EASC. Indeed, when offering it, $S$ must
make sure to guarantee him an expected payoff that is at least as high as the one he
could get if only EPSC were available.

We solve this game by backward induction. First, we compute $S$’s optimal design of
a side-contract EASC, given a grand-contract GC. Secondly, we solve for the optimal
grand-contract GC given $S$’s best reaction.

For a given grand-contract GC, $S$ reacts by designing an EASC. In a sense, $S$ is
renegotiating the terms of the GC offered by $P$ to $A$. Any given EASC induces a choice
by $A$ of either $\bar{c}$ or $\underline{c}$. Thus, there is no loss of generality in focusing only on the $EASC$ that, for a given action, guarantee $S$ the highest expected payoff, given transfers from $P$ and side-transfers that may be exchanged with $A$.

Suppose $S$ wishes to induce $A$ in to choosing action $c'$ instead of $c''$. She then chooses a schedule of side transfers $y_{\pi m(\sigma)}$ and a reporting function $m(\sigma)$ to solve the following program:

$$\max_{\{y_{\pi m(\sigma)}, m(\sigma)\}} \sum_{\sigma} \left( E_{\sigma}^{c'} s_{\pi m(\sigma)} - E_{\sigma}^{c'} K \left( y_{m(\sigma)} \right) y_{m(\sigma)} \right) \quad \text{s.t.}$$

$$\sum_{\sigma} E_{\sigma}^{c'} (t_{m(\sigma)} + y_{m(\sigma)}) - \psi (e') \geq \sum_{\sigma} E_{\sigma}^{c''} (t_{m(\sigma)} + y_{m(\sigma)}) - \psi (e'') \quad \text{(SIC)}$$

$$\sum_{\sigma} E_{\sigma}^{c'} (t_{m(\sigma)} + y_{m(\sigma)}) - \psi (e') \geq u^A \quad \text{(SPC)}$$

$$s_{\pi m(\sigma)} \geq y_{\pi m(\sigma)} \geq -t_{\pi m(\sigma)} \quad \forall \{\pi, m(\sigma)\} \quad \text{(3)}$$

$u^A$ is $A$’s outside option when refusing the $EASC$, i.e. the expected payoff (net of bribes paid to $S$) when side contracting takes place only at Stage 4. $S$ must ensure that $A$ indeed prefers the intended action $c'$ to $c''$ ($\text{(SIC)}$). She also has to make sure that participating in the side-agreement makes $A$ better off ($\text{(SPC)}$) than simply postponing side-contracting to Stage 4. Finally, side-transfers cannot be larger than the payments promised by $P$. It is intuitive that, for any given $GC$, there is always an $EASC$ that induces action $c$ (solution to the above problem). We will denote by $E^c(s - y)$ the expected payoff derived by $S$ when offering the most profitable $EASC$ inducing $c$.

---

18 We do not write it explicitly as it varies non continuously with the payments designed in the $GC$. See the Appendix for a detailed description. For expository clarity we have also omitted constraints ensuring that the reporting function $m(\sigma)$ be interim rational. It is however trivial that $S$ is capable of finding a schedule of out-of-equilibrium transfers ensuring interim rationality of any such reporting function.

19 Since this constraint never binds, the model could easily be extended to consider $\psi$ as the monetary opportunity cost of not breaching the law or accepting bribes.
P’s problem is that of designing a grand-contract $GC$ such that $S$ proposes $A$ an $EASC$ inducing high effort $\bar{e}$, while minimizing expected payments:

$$\min_{\{t = m, s = m \in \pi, m\}} \sum_{\sigma} E^e_{\sigma} \left(s_{m(\sigma)} + t_{m(\sigma)}\right) \quad \text{s.t.}$$

$$E^e(s - y) \geq E^x(s - y) \quad (IC1)$$

That $S$ designs an $EASC$ inducing high effort rather than one inducing shirk is ensured by $[IC1]$. Given the $EASC$ presented above, such constraint makes incentive compatibility and participation constraints at the agent’s level redundant. As long as it holds, $S$ will, if necessary, top up the transfers promised to $A$ by $P$ in order to have him choose $\bar{e}$. However, it is optimal for $P$ to make sure that this never occurs. The reason is that side-transfers involve transaction costs. Therefore, in any state of the world, it is always less costly for $P$ to make a transfer directly to $A$, in order to provide incentives, than transfer money to $S$ and have her pay $A$.

The form of the $EASC$s inducing high and low effort depend on the ordering of state-contingent salaries to $S$ and $A$. This, in turn, determines the form of $[IC1]$. A complete characterization of such ordering and of the side-contracts resulting from it is provided in the Appendix.\[^{20}\] We do not present them here for reasons of brevity, but we provide an informal discussion of their nature.

To begin, suppose $\bar{e}$ was the action $S$ decides to induce. The $EASC$ can then be referred to as a cooperative side-contract. This is because it is in $S$’s own interest to commit to preserve $A$’s incentives to work hard. Importantly, this could not be done if side-contracting happened only at Stage 4, since, by assumption, no prior engagements between $A$ and $S$ can be made. $S$ has to let $A$ pocket enough transfers to compensate\[^{20}\]

\[^{20}\]To give an idea, it is such that salaries paid to $S$ when reporting $B$ or $N$ are bounded from above, in order to avoid the possibility of her using them to compensate $A$ for making low effort and abuse the incentive scheme.
the cost of effort $\psi$, as well as $u^A$. Given the inefficiency of side-transfers with respect to transfers from $P$ (represented by the cost $k > 1$ of transferring money), it is intuitive that the best way to do so is to let $A$ collect payments designed to him by $P$. Thus, $S$ will always report $G$ truthfully. As for $N$, $S$ either reports it truthfully or she reports $G$, depending on how $P$ designs her salary. We show in the Appendix that (under the optimal ordering of state-contingent salaries to $S$) the payoff for her when offering the $EASC$ inducing $\bar{e}$ is:

$$E^\bar{e}(s - y) = \max_{m(N) = G, N} E^G_G s_G + E^N_N s_m(N) + \frac{1}{k} \left( E^G_G t_G + E^N_N t_m(N) - \psi - u^A \right)$$ (4)

Consider now the $EASC$ inducing $e$. It can be referred to as a collusive side-contract. This is because $S$ deliberately induces $A$ into shirk, while committing to cover up the proofs, i.e. to $m(B) = G$. $S$ also commits to $m(N) = N$ or $G$, depending on how $P$ designs her salary. This is the most profitable way for $S$ to induce $A$ into shirk. Indeed, $A$ could be induced into shirk also by letting him perceive, before he chooses his action, a threat of being extorted and deprived of his salary when $G$ turns up. However, precisely because $A$ would shirk, $G$ would not turn up at all. Anticipating this, $S$ is better off choosing to induce shirk by colluding with $A$. Also, $S$ is better off exploiting transfers paid by $P$, rather than transferring money to $A$ directly, due to side transfers’ inefficiency. However, she has to make sure she leaves $A$ transfers at least as high as to match the outside option $u^A$. As we show in the Appendix, the payoff for $S$ when such side-contract is implemented is:

$$E^e(s - y) = \max_{m(N) = G, N} E^G_B s_G + E^N_N s_m(N) + \frac{1}{k} \left( E^G_B t_G + E^N_N t_m(N) - u^A \right)$$ (5)

\[21\] This $EASC$ and the one inducing $e$ are such that $(SPC)$ is binding; this is because $S$ can exploit the message $\sigma$ to punish $A$ for eventual deviations.
The following Proposition describes the optimal GC, result of the maximization problem presented above.

**Proposition 2** Suppose S and A can side-contract both before the latter has chosen his action e and after that (i.e. both at Stage 2 and Stage 4). If the supervisory signal σ is insufficiently precise and transaction costs of organizing side-transfers are low enough (i.e. \( p < \frac{2\rho_{\pi}-1}{\rho_{\pi}} \) and \( k \leq \tilde{k} \equiv 1 + \frac{2\rho_{\pi}-1}{p(1-\rho_{\pi})} - \frac{p}{(1-p)(2\rho_{\pi}-1)} \)) the optimal GC is

\[
II: \quad t_{\pi G} = t_{\pi N} = \frac{\psi}{2\rho_{\pi} - 1 + (1 - \rho_{\pi})p} \quad s_{\pi m} = \frac{1}{k} \left( \frac{\psi}{2\rho_{\pi} - 1} - t_{\pi G} \right) \forall m
\]

and all other transfers are set to zero. The total expected payment for P is therefore

\[
E(z)^{II} = \frac{\rho_{\pi}}{k} \left( \frac{(k-1)\psi}{2\rho_{\pi} - 1 + (1 - \rho_{\pi})p} + \frac{\psi}{2\rho_{\pi} - 1} \right)
\]

When instead σ is sufficiently precise or transaction costs of side-transfers are high (i.e. \( p \geq \frac{2\rho_{\pi}-1}{\rho_{\pi}} \) or \( k > \tilde{k} \)), the optimal GC is

\[
III: \quad t_{\pi G} = \frac{\psi}{p\rho_{\pi}} \quad s_{\pi G} = \max \left( \left( \frac{p}{(2\rho_{\pi} - 1) - \frac{1}{\rho_{\pi}}} \right) \frac{\psi}{pk}, 0 \right) < s_{\pi N} = s_{\pi G} + \frac{t_{\pi G}}{k} \quad s_{\pi B}(0; s_{\pi N})
\]

and all other transfers are set to zero. The total expected payment for P is therefore

\[
E(z)^{III} = \left( \frac{\rho_{\pi}}{2\rho_{\pi} - 1} - 1 \right) \frac{\psi}{k} + \psi
\]

**Proof.** See Appendix. □

**Discussion of the result.** When S is not benevolent and in the absence of ex ante side-contracting, P’s optimal choice is to offer an incentive scheme to A in which N
and $G$ are treated in the same way. In contrast, when the possibility of ex-ante side-
contracting is considered, $P$ can use $GC$ III and offer $A$ a payment schedule as steep as if $S$ were benevolent, in which $t_{\pi G} > t_{\pi N} = t_{\pi B} = 0$. In a nutshell, this is because the scope for extorting $A$ is drastically reduced. If $S$ wanted to design a side contract inducing high effort $\bar{e}$ from $A$, extorting him would then be completely counterproductive. Indeed, if $A$ perceives the possibility of being extorted, he will not work hard.

One may then ask how $P$ can make sure $S$ will choose to offer the $EASC$ inducing $\bar{e}$. Indeed, she may be tempted to do the opposite, for two reasons: first, she may hope to capture the bribes $A$ would be willing to pay in order to have her hide proofs of misbehavior. Second, she may capture the rewards $P$ may pay as "bounties" for gathering and reporting such proofs. We find that, in order to prevent this, $P$ optimally promises $S$ a reward $s_{\pi N}$ for truthfully reporting $N$. Importantly, this payment is conditional on $\pi = \bar{\pi}$. Since high profits are more likely when $A$ works hard, this incentivizes $S$ to preserve $A$’s incentives to make $\bar{e}$. However, this payment may not be enough, in particular when $p$ is sufficiently large (so $N$ turns up too rarely). In that case, it is optimal to reward the supervisor also when reporting $G$. Again, $P$ will optimally make payment $s_{\pi G}$ conditional on $\pi = \bar{\pi}$ as an incentive for $S$ to make sure $A$ chooses high effort.

Our interpretation of $GC$ III is that, when facing the possibility that $S$ actively affects $A$’s incentives, $P$ optimally stimulates her interest in the successful outcome of the project she is supervising. With only ex post side contracting $S$ plays no active role in incentivizing $A$. We found, indeed, that the optimal $GC$ was such that no reward was to be paid to $S$ in equilibrium. When we open up the possibility of ex ante side contracting, the optimal $GC$ becomes such that $S$ is left with a positive expected transfer (an information rent) in equilibrium. This is because, when ex ante

\[ III \text{ is such that } s_{\pi G} \in [0; s_{\pi N}) \text{ and } s_{\pi N} > s_{\pi G}. \] This means that $S$ can, a priori, threaten to report either $N$ or $B$ when $G$ turns up. However, she will find it optimal to commit ex-ante not to engage in such behavior, for the reasons explained above.

22
side contracting is allowed, collusion is not only about protecting rents, but also about increasing them.

GC III is however not without drawbacks. First, if $S$’s information is not very precise, i.e. if $p$ is low, payment $s_{\#N}$ is made often and its value increases (along with $t_{\#G}$). Second, if $k$ is relatively small, the probability of collusion to $S$ is high, which raises the size of $s_{\#N}$ needed to prevent it. As a consequence, as long as side contracting is possible at both Stage 2 and 4, GC III is optimal if and only if either $p$ or $k$ are sufficiently high (see the thresholds provided in Proposition 2). If not, the optimal GC is II. Such contract is quite similar to GC I derived in the presence of only ex post side-contracting: $P$ gives up on distinguishing payments when either $G$ or $N$ are reported. Therefore, $t_{\#G} = t_{\#N}$ and $s_{\#G} = s_{\#N}$. The price to pay for doing so is, as explained above, that $S$’s information is used only to a partial extent. Importantly, GC II must be such that $S$ is promised rewards when $\pi = \bar{\pi}$. The reason is, again, to avoid the possibility of her inducing $A$ into making low effort. Observe that this implies that GC II necessarily involves strictly higher payments for $P$ with respect to GC I. This has implications for the optimal timing of supervisory activities, as we show in the following section.

3.4 Should supervisors perform monitoring or auditing functions?

We now compare the optimal $GC$ obtained in the presence of ex ante side contracting (i.e. allowing $A$ and $S$ to side contract starting from Stage 2 and through Stage 4 of the game) and in its absence (if Stage 2 is void). The objective is to identify conditions under which allowing for ex ante side contracting is beneficial for principal. This has implications for the optimal organizational design in terms of the timing of supervisory activities. In particular, we interpret this result as an indication of the optimality
of having $S$ participate in the organization from the start of the project as opposed to having him check $A$’s behavior only after the latter has made effort. Borrowing terminology used in Strausz (2006), this distinction can be recast in terms of designing an organization in which the supervisor functions as a monitor rather than one in which he functions as an auditor.\footnote{Strausz (2006, p.1) states: “The difference between the two procedures is that monitoring takes place while the agent chooses his action whereas auditing occurs after he has taken his action”. However, differently from his paper, in our setup auditing takes place before any additional signal about the agent’s behavior is available.} Proposition 3 illustrates the results.

**Proposition 3** There exists a threshold $k_x \geq \tilde{k}$ such that when the transaction cost of side contracting $k$ is sufficiently high, i.e. $k > k_x$, $P$ is better off letting $S$ and $A$ side-contract both at the ex ante and at the ex post stage. $S$ should then perform monitoring functions. Otherwise, $S$ should be allowed to side contract with $A$ only at the ex post stage and perform only auditing functions.

**Proof.** See Appendix \footnote{It is straightforward to see that $GC II$ can never be preferrable to $GC I$. Therefore, in all circumstances in which $GC II \succ GC III$ appointing $S$ at the later stage is optimal.}

**Intuition.** Proposition 2 suggests that, under ex-ante side-contracting, if $k > \tilde{k}$, $P$ offers $A$ the same payment scheme as in the “benevolent supervisor” benchmark. If Stage 2 is void (as in Proposition 1) a flatter payment scheme has to be offered to prevent such opportunistic behavior. $P$ may therefore gain from having $S$ take part in the organization from the outset, as long as $A$ can be left with a lower rent. However, under ex ante side contracting a rent has to be left to $S$. This is to prevent her from designing a collusive $EASC$. The principal thus faces a trade-off when deciding when to appoint $S$. Proposition 3 indicates such trade-off favors early appointment of $S$ whenever transaction costs of side-contracting are sufficiently large. This is because larger transaction costs reduce the profitability of collusion. Then, the gains from reducing the rent left to $A$ in $GC I$ outweigh the increased rent left to $S$ in $GC III$.\footnote{It is straightforward to see that $GC II$ can never be preferrable to $GC I$. Therefore, in all circumstances in which $GC II \succ GC III$ appointing $S$ at the later stage is optimal.}
This result runs somewhat against the common wisdom that auditing tasks should either be separated amongst distinct entities or at least delegated to supervisors appointed only for short periods of time. Such an organization of auditing activities is thought to make collusion more difficult, chiefly because supervisor and supervisee do not have time to develop the kind of relationships with trust or a sentiment of reciprocity. What our result suggests is that by reducing the factors that are thought to facilitate collusion one necessarily also reduces the factors facilitating cooperation. In particular, in our model, by having the supervisor carry only ex-post audits, the principal foregoes the opportunity to make him an essential input in the production process. Whatever the behavior of the supervisor, the agent has already chosen his action and thus the realization of profits may no longer be altered. On the contrary, if appointed early, the supervisor can be made to internalize the consequences of its behavior, and in particular of extortion. In Hiriart, Martimort and Pouyet (2011), for instance, two rounds of collusion are also present: before and after public information realizes (but both rounds after the action). They conclude that is better to have separate entities carry the two audits so as to reduce the scope for ex-ante collusion. Although we ask ourselves a different question, our model suggests that it may be better to appoint the supervisor ex-ante to reduce the scope for ex-post side-contracting.

3.5 Delegation

Note also that this result is in contrast to the verifiable information case. In such a case the principal would make a payment to the agent only when proofs of good behavior are transmitted (see Kessler [2000]). In such organizations it is easy to show that the firm is indifferent between delegating payroll authority or not.
4 Augmented Mechanisms

The analysis so far has arbitrarily restricted the scope for communication between the principal and the other layers of the hierarchy. In particular, the grand contract \( GC \) presented in the previous section was computed under the assumption that the principal could communicate exclusively with the supervisor, and only concerning the realization of the latter’s signal. This, while reasonable and relevant for many applications in mind, nevertheless gives rise to some concerns on theoretical grounds. In this section we adopt a more mechanism design approach to the problem at hand and relax all restrictions concerning what the principal can do. As Maskin and Tirole (1999) argue, we take the view that no mechanism consistent with the set of assumptions should be disregarded a priori. To see why the grand contract presented in the previous section might have been restrictive note that supervisor and agent each possess private information that might be of interest to the principal. The agent, for instance, knows the realization of \( \sigma \). The principal, in an effort to combat opportunism by the supervisor, could use such correlated information to punish conflicting reports. In the same spirit, so as to deter collusion, the principal could ask, and reward, an agent revealing the fraudulent scheme. In this spirit, we hereafter abusively denote \( i_A \in I_A \) and \( i_S \in I_S \) the private information of the agent and the supervisor, where \( I_S \) and \( I_A \) denote all possible information sets that can be reached by the supervisor and the agent in the game. As an example note that both \( i_A \) and \( i_S \) include \( \sigma \) (once it has realized), as nested information is here assumed. Similarly, if an EASC has been offered by the supervisor to the agent then both \( i_A \) and \( i_S \) include EASC, as well as the acceptance, or not, of \( A \). This last piece of private information will is of crucial importance in the design of the optimal grand contract \( GC \): the principal is able to elicit EASC and punish a supervisor behaving opportunistically.

We now present both grand and side contracts, as well as their associated optimization problems, and highlight how these differ from those considered in the previous
section.

**Grand contract.** A grand contract \( GC \) is a sequence of \( N \) communication rounds and a collection of two arbitrary message spaces, \( M_S \) and \( M_A \), and profits \( \pi \), as well as two functions defined on the product of these spaces. These two functions specify, respectively:

1. The transfer to \( A \), \( t: \pi \times M_S \times M_A \rightarrow \mathbb{R}_+ \),
2. The transfer to \( S \), \( s: \pi \times M_S \times M_A \rightarrow \mathbb{R}_+ \).

We denote \( m_A \) and \( m_S \) generic elements belonging to, respectively, \( M_A \) and \( M_S \). While no restrictions are imposed on the timing of communication between players we do not make this explicit in our notation. One should thus consider \( m_j \), for \( j = S, A \), as specifying a message to send at each round of communication \( n \), for \( n = 1, \ldots, N \). Compared to the grand contract developed in the previous sections, the principal now (potentially) communicates not only with the supervisor but also with the agent. No restrictions are imposed on what this communication may be about. In addition, and importantly, the principal may communicate at any point in time during the game, i.e. not only once \( \sigma \) has realized.

Since the game is solved by backward induction, we present the principal’s optimization problem after having discussed the supervisor’s problem(s).

**Side contract.** Let us first comment on the ex ante side contract. \( EASC \) is a mapping from profits \( \pi \), \( S \)'s private information \( I_S \), the collection of message spaces \( M_S \), \( M_A \) and \( R_A \), where \( R_A \) denotes the space of possible internal reports (from \( A \) to \( S \)), to three functions defined on the product of these spaces. We denote \( r_A \) the internal reporting strategy of \( A \) to \( S \). The three functions specify:

1. The side transfer, \( y: \pi \times I_S \times M_S \times M_A \times R_A \rightarrow \mathbb{R} \),
2. S’s reporting strategy, $m_S$: $\pi \times I_S \times M_S \times M_A \times R_A \rightarrow M_S$.

3. A’s reporting strategy, $m_A$: $\pi \times I_S \times M_S \times M_A \times R_A \rightarrow M_A$.

The Revelation Principle holds at the side contracting stage, where $M_A \times e$ is A’s strategy space. It is thus without loss of generality for the supervisor to design a direct mechanism $EASC$ such that the agent internally reveals its private information, i.e. such that $r_A^* = i_A$. From the principal’s perspective, therefore, supervisor and agent share the same private information. This remark suggests that the $P$ can potentially elicit the same information from either players. Combined with the fact that, by an appropriate choice of out-of-equilibrium side transfers, $S$ and $A$ can costlessly coordinate their reports into $GC$, this will allow us to establish that it is without loss of generality for $P$ to restrict communication with only one layer of the hierarchy. Compared to the side contract computed in the previous sections, we now also remove restrictions on the private communication between $S$ and $A$. For convenience, we assume that side transfers no longer involve any transaction costs, i.e. $k = 1$. This will only reinforce our results. Suppose $S$ wishes to induce action $e'$. Her problem then takes the form:

$$\max_{\{y_{mr}, m\}} \left\{ \sum_{\sigma} E_{\sigma}^{e'} (s_m - y_{mr}) \right\} \quad \text{s.t.}$$

$$\sum_{\sigma} E_{\sigma}^{e'} (t_m + y_{mr}) - \psi(e') \geq \sum_{\sigma} E_{\sigma}^{e''} (t_m + y_{mr}) - \psi(e'') \quad (SIC)$$

$$\sum_{\sigma} E_{\sigma}^{e'} (t_m + y_{mr}) - \psi(e') \geq u_{e'}^A \quad (SPC)$$

$$s_{\pi m} \geq y_{\pi mr} \geq -t_{\pi m} \quad \forall \{\pi, m, r\} \quad (SLL)$$

where $m$ now denotes a vector of reporting strategies $(m_A, m_S)$ specifying messages to be sent into $GC$ at each round of communication $n$. As before, $S$ maximizes his payoff subject to the agent choosing the desired action $(SIC)$, wishing to participate $(SPC)$, and limited liability constraints $(SLL)$. In addition to choosing side transfers
(which are also contingent on internal reports $r$) and own reporting strategy $m_S$, the supervisor now also chooses the agent’s reporting strategy $m_A$ into GC. This is achieved costlessly by choosing an appropriate set of side transfers.

Note that $S$, if he wishes, may postpone side contracting until the ex post stage, i.e. once $\sigma$ has realized. $EPSC$ then resembles $EASC$ in every respect except that (i) $e$ can no longer be influenced, (ii) $\sigma$ has realized and (iii) message spaces need to be modified to account for the fact that some communication stages may already have taken place. To continue let us briefly comment on the $A$’s reservation utility $u^A_e$. Inequality (SPC) tells us that $A$ should be better off participating in the $EASC$ than refusing it and obtaining utility payoff $u^A_e$. This reservation utility $u^A_e$ represents the payoff to $A$ when playing non cooperatively with $S$ until at least the ex post side contracting stage (assuming the game proceeds until then). Note that $u^A_e$ is contingent on the action $e'$ the rejected $EASC$ was attempting to induce. This is in contrast to the side-contract considered in previous section, where it was instead the case that $u^A_e = u^A_e = u^A$.

By communicating more extensively with the hierarchy $P$ may be able to elicit the nature of the (rejected) $EASC$ and thus make reservation payoffs contingent on this information.

**Principal’s problem.** Proceeding by backward induction, $P$’s problem now takes the form:

$$\min \left\{ \sum_{\sigma} E^p_{\sigma} (s_{m} + t_{m}) \right\} \quad \text{s.t.}$$

$$\sum_{\sigma} E^c_{\sigma} (t_{m} + y_{m}) - \psi \geq \sum_{\sigma} E^c_{\sigma} (t_{m} + y_{m}) \quad (6)$$

$$\sum_{\sigma} E^c_{\sigma} (s_{m} + y_{m}) \geq \sum_{\sigma} E^c_{\sigma} (s_{m} + y_{m}) \quad (7)$$

In words, $P$ minimizes expected salary costs subject to $S$ preferring to design an
EASC inducing $e = \bar{e}$ than $e = e$ (7), and subject to $A$ preferring to exert high rather than low effort (6). Note that (6) is superfluous in that it is implied by (7); its statement will however turn out to be useful later.

4.1 Solving for the optimal grand contract

Prima facie, the computational difficulty associated with designing a grand contract $GC$ contingent on the full private information of supervisor ($i_S$) and agent ($i_A$) is overwhelming. In this section, however, we show that while $P$ can indeed gain by communicating more extensively with her hierarchy, the optimal contract takes a rather simple form. We show that by including a self reporting scheme in $GC$, ex ante side contracting brings no additional costs compared to ex post side contracting only, even in the absence of transaction costs.

Self reporting scheme. We say that $GC$ includes a self reporting scheme if it specifies a communication stage after $S$ has (possibly) designed EASC and before $A$ has chosen $e$. In this communication stage, $S$ is asked to report on the status of ex ante side negotiations.

Recall that we denote $u^A$ and $u^S$ the payoffs to, respectively, $A$ and $S$ when playing non cooperatively the ex ante side contracting stage. In addition, we denote $u^A_{\bar{e}}$ and $u^A_{\tau}$ the reservation payoff of $A$ when offered by $S$, respectively, an opportunistic and a cooperative ex ante side contract. Note that in the previous section it was the case that $u^A_{\bar{e}} = u^A_{\tau} = u^A$. The next lemma gives a preview of the scope of self reporting scheme.

Lemma 1 It is without loss of generality for the principal to restrict his attention to grand contracts including a self reporting stage. Moreover, the principal finds it optimal to set reservation payoffs $u^A_{\bar{e}}$ to zero and $u^A_{\tau}$ arbitrarily large.

\footnote{Although it should be noted that the sequential nature of our game annihilates the infinite regress problem present in most common agency models.}
Proof. See Appendix ??.

Intuition for optimality of self reporting schemes. In a nutshell, it is optimal to include a self reporting scheme in the grand-contract as one can then only enlarge what $P$ can do. By specifying a communication round after $S$ possibly designs an $EASC$, and before $A$ chooses his effort level $e$, $P$ is able to elicit, to a large extent, the status of ex ante side negotiations. At the self reporting stage $S$ is given three payoff equivalent choices. The supervisor is asked (a) whether the agent rejected an opportunistic $EASC$ ($message \ a$), (b) whether the agent rejected a cooperative $EASC$ ($message \ b$) and (c) whether either an $EASC$ was agreed upon or no $EASC$ was ever designed ($message \ c$). 

If $S$ sends message $a$, $P$ then makes payments $t = u^A_e$ and $s = u^S$ and the game ends. Similarly, if $S$ sends message $b$, $P$ makes payments $t = u^A_e$ and $s = u^S$ and the game ends. Finally, if $S$ sends message $c$ the game proceeds and expected payoffs from the ex post side contracting stage are $u^S$ and $u^A$. The supervisor, being indifferent, acts truthfully, thereby revealing the status of ex ante side negotiations. Note, importantly, that $u^A_e$ and $u^A_e$ are then $A$’s reservation payoffs when offered, respectively, a cooperative and an opportunistic $EASC$. Since $P$ can always set $u^A_e = u^A_e = u^A$ it follows that any equilibrium induced by a grand-contract not including a self reporting stage may be replicated by a grand contract including one.\footnote{The effectiveness of the described self reporting scheme relies on two assumptions. First, the way in which the principal is capable of eliciting the supervisor’s private information relies importantly on the standard, yet strong, assumption whereby an agent indifferent between several strategies selects the one most preferred by the principal. Second, the commitment to stop the game before the agent has even chosen his action is also strong and may not be interim rational.}

Choice of reservation payoffs. By including a self reporting scheme in its grand-contract $P$ is able to treat $u^A_e$ and $u^A_e$ as distinct choice variables. This is a very powerful tool when bearing in mind that reservation payoffs determine $A$’s bargaining strength vis-à-vis $S$, and thus ultimately actual payoffs. While reservation payoffs were also determined by $P$ in the previous sections, these were not contingent on the nature of
the EASC. To minimize the sum of expected payments it is optimal for $P$ to relax the incentive compatibility constraint at $S$’s level $[7]$. Ideally, therefore, $P$ would like to both relax the agency problem faced by a “cooperative” supervisor and exacerbate that of an “opportunistic” supervisor. Intuitively, this is reached by setting $u^A_e$ to zero and $u^A_e$ large enough.

Lemma 2 It is without loss of generality for the principal to communicate exclusively with the supervisor and restrict communication to (i) the status of ex ante side negotiations and (ii) the realization of signal $\sigma$.

Proof. See Appendix ??.

Intuition. Through its design of an ex ante side-contract, by the Revelation Principle, $S$ finds it optimal to elicit all of $A$’s private information $i_A$. Supervisor and agent’s private information, from the perspective of the principal, are thus identical. It follows that all the information that $P$ can elicit from $A$ can possibly also be elicited from $S$. In addition, with an appropriate choice of side transfers, $S$ and $A$ are able also to costlessly coordinate their reports into the grand-contract. In particular, $S$ being given full bargaining power, $A$’s reporting strategy is essentially chosen by the former.

In a sense, therefore, communication is de facto occurring only with one player and it is therefore not surprising that the principal does not gain from communicating also with $A$. Finally, by eliciting from $S$ the realization of $\sigma$, and by making use also of the profits $\pi$, the principal has enough “degrees of freedom” to replicate any expected transfers that would arise under broader communication games.

The two previous lemmas imply that although there is scope for more extensive communication, namely by including a self reporting scheme, such communication is still of limited nature, and thus complexity. The following proposition, building on these results, states the optimal mechanism. Let us recall the notation. We first review
the messages at the disposal of $S$ during the self reporting scheme. Message $a$ is meant to indicate that $A$ rejected an opportunistic ex ante side contract. Similarly, message $b$ is meant to indicate that the agent rejected a cooperative ex ante side-contract. Finally, message $c$ means that either (i) the coalition agreed on an ex ante side-contract or (ii) that $S$ did not make an offer. The supervisor, as is suggested by lemma 2, is also asked a report on the realization of signal $\sigma$. When $S$ sends message $G$ (resp. $B$) she is claiming to have observed $A$ working hard (shirking). A supervisor sending message $N$ indicates that she is unaware of the effort level chosen by the agent.

**Proposition 5** The optimal grand contract is such that $M_A = \emptyset$ and $M_S = \{\{a, b, c\}, \{G, B, N\}\}$. The supervisor sends one first message in the set $\{a, b, c\}$ after the ex ante side contracting stage and a second message in the set $\{G, B, N\}$ once $\sigma$ has realized. Payments are such that

$$t_{\sigma G oc} = t_{\sigma N oc} = \frac{\psi}{2\rho_\pi - 1 + (1 - \rho_\pi) \rho_\sigma},$$

and all other transfers to zero. The expected payroll cost thus is

$$E(z)^I = \frac{\rho_\pi \psi}{2\rho_\pi - 1 + (1 - \rho_\pi) \rho_\sigma},$$

i.e. identical to that under only ex post opportunism.

**Proof.** See Appendix ??.

**Comments and intuition.** What Proposition 5 says is that (i) despite side contracting occurring ex ante and (ii) despite the absence of transaction costs $k$, the principal’s
payoff is identical to that under only ex post opportunism. Recall that in the previous section, Proposition 4 established that ex ante opportunism was more harmful than the ex post one precisely when \( k \) was low. The main reason behind \( P \)'s ability to reduce drastically the costs of side contracting, compared to the analysis in the previous sections, resides in the usefulness of self reporting schemes, i.e. resides in the gains reaped from communicating more extensively with the hierarchy. In particular, as highlighted by lemma 1, \( P \) is capable of setting distinct \( A \)'s reservation payoffs according to the behavior of \( S \) (opportunistic or cooperative). More precisely, \( P \) drives the payoff of an "opportunistic" \( S \) to zero by setting \( u^A_S \) large enough. By doing so \( P \) ensures that it is not in \( S \)'s best interest to let shirk happen. Said differently, the biggest threat created by ex ante side contracting is removed.\(^{27}\) Information manipulation, however, remains a problem, even in the presence of a cooperative side-contract. Indeed, even if \( S \) wishes \( A \) to work hard, the temptation to inflate and share the latter's salary is still relevant. Not surprisingly, therefore, the optimal contract collapses to the optimal contract we computed in the presence of ex post opportunism only.

**Proposition 6**  
*An arm's length relationship is strictly suboptimal.*

**Proof.** See proof in Appendix. ■

**Intuition.** Delegating the task of contracting with the agent to the supervisor, while keeping open the communication channel with the agent through the design of a whistle-blowing program, is *strictly* optimal. By creating an *arm's length relationship* the principal is able both to indirectly enjoy the supervisor’s superior information and make payments to the latter not responsive to reports, thereby eliminating any scope for opportunism.

\(^{27}\) One could then think that \( S \) may wish to coordinate \( m_A \) and \( m_S \) so as to pocket \( u^A_S \). However, by definition of what a reservation payoff is, \( S \)'s payoff would then be identical to that in case the game proceeds to the ex post side contracting stage. Under the assumption that, when indifferent, \( S \) does what \( P \) would like him to do, this strategy is thus irrelevant.
As in the benevolent supervisor benchmark, the supervisor’s private information is such that the agent is offered the first best contract. No rent is given up to the agent. The principal, anticipating this, must then find the cheapest way to reimburse the supervisor’s (expected) salary costs $\psi$. Recall that through the whistleblowing program, the supervisor’s payoff when inducing shirk is set to zero. In other words, the supervisor’s payoff when deviating is equal to zero. The principal then simply rewards the supervisor when profits are high by making positive payment $\frac{\psi}{\rho_n}$, thereby also leaving no rent to the supervisor, while reimbursing him for his salary expenses.

No centralized contracting scheme can replicate this outcome. In a centralized setting, to have the supervisor be useful, transfers to the agent must be, at least partially, contingent on the former’s reports. This however, almost mechanically, implies introducing some (costly) scope for information manipulation, either by the coalition or the supervisor alone. This is shown in the Appendix, in which no restrictions were imposed on what the principal could do under centralized contracting.

We have thus established that by the use of augmented mechanisms side-contracting does not bring any additional cost to the hierarchy, despite the absence of transaction costs or asymmetric information at the coalitional level. Only the decentralized organization can achieve this outcome. Despite the use of similar, at least in spirit, whistleblowing programs Celik (2009) concludes that delegation is usually dominated in adverse selection set ups. In a moral hazard set up, but for a different hierarchy, Baliga and Sjostrom (1998) show that decentralization is optimal. Itoh (), in once again a moral hazard set up, in constrast to us, argues that message games do not help the principal.
5 Concluding remarks

We have investigated the impact of informal agreements on the functioning of organizations, using a model based on a three-tier-hierarchy framework (in the spirit of Tirole’s (1986) seminal paper), with a moral hazard problem at the bottom and soft supervisory information. We have highlighted the importance of the timing of appointment of supervisors to contrast the pervasiveness of opportunistic behavior. In particular, our results suggest that allowing supervisor and supervisee(s) to side-contract before the latter chooses his action can be beneficial to the organization. The optimal incentive scheme provides group-based incentives, thus contrasting collective opportunism, while also eradicating individual opportunism. This is because if the supervisor is interested in the positive outcome of the agent’s task, then she has no incentives to act opportunistically vis-à-vis the latter. We consider delegation and whistleblowing...

Our results rest on some important assumptions...
Appendix

6 Proof Proposition 1

Throughout the proof we make the following assumption:

**Assumption 1** \( t_{\pi G} \geq t_{\pi N} \geq t_{\pi B} = 0, \ \forall \pi \in \{\bar{\pi}, \bar{\pi}\} \)

Assumption 1 is w.l.o.g and simplifies the exposition of the proof. We proceed by first solving for the optimal \textit{EPSC} for a given \textit{GC}, and then moving to the principal’s problem.

6.1 Ex Post Side Contracting Stage

**Notation.** Denote \( v_{\pi \sigma} \) \( S \)'s posterior belief over the distribution of \( \pi \), upon having observed the realization of \( \sigma \). Note that observing \( \sigma = G, B \) is perfectly informative of \( e \), so \( v_{\pi G} = p_{\pi G}^e \) and \( v_{\pi B} = p_{\pi B}^e \). When \( \sigma = N \), we assume \( v_{\pi \sigma} \) is computed anticipating that \textit{GC} gives \textit{A} sufficient incentives to make high effort, so \( v_{\pi N} = p_{\pi N}^e \). In equilibrium this is confirmed.

**Supervisor’s problem.** Suppose no \textit{EASC} was agreed upon at Stage 2. At Stage 4, \( S \)'s objective is then to maximize

\[
\sum_{\pi} v_{\pi \sigma} \left( s_{\pi \sigma} - \frac{1}{k} \cdot y_{\pi \sigma} \right)
\]

i.e. the sum of the expected transfers from \textit{P} and \textit{A}. To do so, \( S \) chooses the report \( m \), where \( m = G, N, B \), to make to \( P \), as well as an out-of-equilibrium report \( l \) (a threat), where \( l = G, N, B \), in case \( A \) refuses \textit{EPSC}. When choosing \( l \), \( S \) needs to ensure that the threat is credible. To illustrate this assume that a given realization of \( \sigma \) has occurred. \( S \) can then credibly threaten \( A \) with a report of \( l \) only if either \( l \)...
is the report that strictly maximizes the expected payment $\sum_{\pi} v_{\pi \sigma} s_{\pi l}$ or $l = \sigma$. More precisely, $S$ may set $l = B$ if (i) $\sum_{\pi} v_{\pi \sigma} s_{\pi B} > \sum_{\pi} v_{\pi \sigma} s_{\pi x}, \forall x = G, N$ and also if (ii) either 

$\{B, N\} = \text{arg max}_x \left( \sum_{\pi} v_{\pi \sigma} s_{\pi x} \right)$ or $\{B, G\} = \text{arg max}_x \left( \sum_{\pi} v_{\pi \sigma} s_{\pi x} \right)$, but $\sigma = B$. In words, if (i) then reporting $B$ is strictly optimal and thus the threat is credible. If (ii), while there are other reports as lucrative as (but no more than) $B$, the supervisor, by assumption, weakly prefers being truthful, thereby making the threat credible.

Similarly, $l = N$ is a credible threat if $\sum_{\pi} v_{\pi \sigma} s_{\pi N} > \sum_{\pi} v_{\pi \sigma} s_{\pi x}, \forall x = G, B$. Also, if either $\{N, B\} = \text{arg max}_x \left( \sum_{\pi} v_{\pi \sigma} s_{\pi x} \right)$ or $\{N, G\} = \text{arg max}_x \left( \sum_{\pi} v_{\pi \sigma} s_{\pi x} \right)$ then $l = N$ is a credible threat if and only if $\sigma = N$. Finally, we have $y_{\pi \sigma m} = 0$ if no viable threats are credible.

**Optimal EPSC.** We illustrate in the following table the EPSC that solves $S$’s problem, for each possible ordering of the transfers specified in $GC$. We denote by $E^e_{\sigma} y_{m(\sigma)}$ the expected side-transfer, as a function of the supervisor’s reporting strategy $m(\sigma)$. We define condition $a$ as $E^e_{\sigma} s_N \geq E^e_{\sigma} \left( s_G + \frac{1}{k} (t_G - t_N) \right)$ and $\bar{a}$ its contrary. Similarly, we define condition $b$ as $E^e_{\sigma} s_B \geq E^e_{\sigma} \left( s_G + \frac{1}{k} t_G \right)$ and $\bar{b}$ its contrary.

---

28 For brevity, we mention only state-contingent side-transfers that would take place in the equilibrium intended by the given EPSC, all the other (out of equilibrium) ones being set to zero.
<table>
<thead>
<tr>
<th>Transfers to S</th>
<th>Transfers to A</th>
<th>$m(\sigma)$</th>
<th>$E^e_{\sigma} y_{m(\sigma)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E^e_{\sigma} s_G = E^e_{\sigma} s_B = E^e_{\sigma} s_N$</td>
<td>$E^e_{\sigma} t_G = E^e_{\sigma} t_B = E^e_{\sigma} t_N$</td>
<td>$\sigma$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$E^e_{\sigma} t_G &gt; E^e_{\sigma} t_B \geq E^e_{\sigma} t_N$</td>
<td>$G$</td>
<td>$E^e_{\sigma} (t\sigma - t_G)$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_G = E^e_{\sigma} s_N &gt; E^e_{\sigma} s_B$</td>
<td>$E^e_{\sigma} t_G = E^e_{\sigma} t_N \geq E^e_{\sigma} t_B$</td>
<td>$G$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$E^e_{\sigma} t_G &gt; E^e_{\sigma} t_N \geq E^e_{\sigma} t_B$</td>
<td>$G$</td>
<td>$0$ if $\sigma = G$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$E^e_{\sigma} (t\sigma - t_G)$ if $\sigma = {N, B}$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_G = E^e_{\sigma} s_B &gt; E^e_{\sigma} s_N$</td>
<td></td>
<td>$G$</td>
<td>0 if $\sigma = G$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-E^e_{\sigma} t_G$ if $\sigma = {N, B}$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_G &gt; \max(E^e_{\sigma} s_B, E^e_{\sigma} s_N)$</td>
<td></td>
<td>$G$</td>
<td>$E^e_{\sigma} (t_N - t_G)$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_N &gt; \max(E^e_{\sigma} s_G, E^e_{\sigma} s_B)$ &amp; $a$</td>
<td></td>
<td>$N$</td>
<td>0</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_N &gt; \max(E^e_{\sigma} s_G, E^e_{\sigma} s_B)$ &amp; $\tilde{a}$</td>
<td></td>
<td>$G$</td>
<td>$E^e_{\sigma} (t_N - t_G)$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_B &gt; \max(E^e_{\sigma} s_G, E^e_{\sigma} s_N)$ &amp; $b$</td>
<td></td>
<td>$B$</td>
<td>0</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_B &gt; \max(E^e_{\sigma} s_G, E^e_{\sigma} s_N)$ &amp; $\tilde{b}$</td>
<td></td>
<td>$G$</td>
<td>$-E^e_{\sigma} t_G$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_B = E^e_{\sigma} s_N &gt; E^e_{\sigma} s_G$ &amp; $b$</td>
<td></td>
<td>$B$</td>
<td>0</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_B = E^e_{\sigma} s_N &gt; E^e_{\sigma} s_G$ &amp; $a$ &amp; $\tilde{b}$</td>
<td></td>
<td>$G$</td>
<td>$-E^e_{\sigma} t_G$ if $\sigma = {G, B}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$0$ if $\sigma = N$</td>
</tr>
<tr>
<td>$E^e_{\sigma} s_B = E^e_{\sigma} s_N &gt; E^e_{\sigma} s_G$ &amp; $\tilde{a}$</td>
<td></td>
<td>$G$</td>
<td>$-E^e_{\sigma} t_G$ if $\sigma = {G, B}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$E^e_{\sigma} (t_N - t_G)$ if $\sigma = N$</td>
</tr>
</tbody>
</table>

### 6.2 Grand Contract

The description of $P$’s problem in designing the GC is provided in the main text. By looking at the table above, one sees that having $E^e_{\sigma} s_G > \max(E^e_{\sigma} s_B, E^e_{\sigma} s_N)$, $E^e_{\sigma} s_N > \max(E^e_{\sigma} s_B, E^e_{\sigma} s_G)$ or $E^e_{\sigma} s_B > \max(E^e_{\sigma} s_G, E^e_{\sigma} s_N)$, would have the agent receive the same net payment, regardless of the true realization of $\sigma$. This cannot be optimal, as it would make $S$’s presence useless. The direct consequence of this is that $P$ has to set $E^e_{BG} s_G = E^e_{BG} s_N = E^e_{BG} s_B$ if it has any hope of making use of $S$’s information. In order
to minimize payments to $S$, $P$ can simply choose $E_G^e s_G = E_G^e s_N = E_N^e s_B = 0$. This implies that

$$\sum_{\sigma} E_{\sigma}^e (t_{m(\sigma)} + y_{m(\sigma)}) - \psi = \sum_{\sigma} E_{\sigma}^e t_\sigma - \psi$$

$$\sum_{\sigma} E_{\sigma}^e (t_{m(\sigma)} + y_{m(\sigma)}) = E_N^e t_N$$

$P$'s problem thus takes the form described in the main text. ■

7 Proof Proposition 2

Throughout the proof we work under Assumption 1. Once again, this is w.l.o.g..

7.1 Proof Lemmas 1 and 2

The following lemma states the five possibly relevant outside options $u^A$ for the agent, in case side-contracting fails at Stage 2. These are computed anticipating the outcome of side-contracting at Stage 4, for a given $GC$.

**Lemma 1** Suppose side-contracting takes place only at Stage 4, then $A$’s expected payoff is:

**Case 1:** $u^A = \max\left(\sum_{\sigma} E_{\sigma}^e t_{\sigma} - \psi, E_N^e t_N\right)$ if $E_G^e s_G = E_G^e s_N = E_N^e s_B$

**Case 2:** $u^A = \sum_{\sigma} E_{\sigma}^e t_N$ if $E_G^e s_G = E_G^e s_N > E_N^e s_B$ or $E_G^e s_N > \max(E_G^e s_G, E_N^e s_B)$

**Case 3:** $u^A = E_N^e t_N$ if $E_N^e s_B = E_G^e s_N > E_G^e s_B$

**Case 4:** $u^A = 0$ if $E_G^e s_G = E_G^e s_B > E_N^e s_N$ or $E_G^e s_B > \max(E_G^e s_G, E_N^e s_B)$

**Case 5:** $u^A = \sum_{\sigma} E_{\sigma}^e t_G$ if $E_G^e s_G > \max(E_G^e s_B, E_N^e s_N)$

**Proof.** This follows directly from Proposition 1. $A$’s payoff is the maximum between the left and the right hand sides of constraint (2). Note that we can be sure that

\[ \text{We should have distinguished, in Case 2, if } E_G^e s_G = E_G^e s_N, \text{ between instances in which } E_G^e t_G = E_G^e t_N \geq E_G^e t_B \text{ and those in which } E_G^e t_G > E_G^e t_N \geq E_G^e t_B. \text{ However, the latter will never take place.} \]
\[ \sum_{\sigma} E_{\sigma}^{e} t_{N} - \psi \leq \sum_{\sigma} E_{\sigma}^{e} t_{N} \text{ in Case 2}, \quad E_{N}^{e} t_{N} - \psi \leq E_{N}^{e} t_{N} \text{ in Case 3}, \quad E_{G}^{e} t_{G} - \psi \leq 0 \text{ in Case 4} \quad \text{and} \quad \sum_{\sigma} E_{\sigma}^{e} t_{G} - \psi \leq \sum_{\sigma} E_{\sigma}^{e} t_{G} \text{ in Case 5}. \]

If these inequalities were violated, then expected payments \( E(z) \) by \( P \) would be larger than \( E(z)^{SB} \).

Throughout the remainder of the proof we will refer to these five cases when we need to substitute for \( u^{A} \). Lemma 2 proves that our results would hold in a set-up in which \( S \) and \( A \) are allowed to collectively renegotiate \( EASC \) at Stage 4.

**Lemma 2** Under collective renegotiation, there is no loss of generality in focusing attention only on \( EASC \).

**Proof.** Consider an \( EASC \) inducing a given action \( e' \), so that

\[ \sum_{\sigma} E_{\sigma}^{e'} (t_{m(\sigma)} + y_{m(\sigma)}) - \psi (e') \geq \sum_{\sigma} E_{\sigma}^{e''} (t_{m(\sigma)} + y_{m(\sigma)}) - \psi (e'') \]

Suppose that, once \( e' \) has been chosen and \( \sigma \) realized, \( S \) decides to renegotiate the \( EASC \). She proposes \( A \) a reporting strategy \( m'(\sigma) \) and a (expected) side-transfer \( E_{\sigma}^{e'} y_{m'(\sigma)} \). Under collective renegotiation, it has to be the case that \( E_{\sigma}^{e'} (t_{m'(\sigma)} + y_{m'(\sigma)}) \geq E_{\sigma}^{e'} (t_{m(\sigma)} + y_{m(\sigma)}) \), for any \( \sigma \). Therefore, \( m'(\sigma) \) has to be such that \( E_{\sigma}^{e'} (t_{m'(\sigma)} + s_{m'(\sigma)}) > E_{\sigma}^{e'} (t_{m(\sigma)} + s_{m(\sigma)}) \): renegotiation must bring to a Pareto improvement for the coalition. However, there is no reason why the \( EASC \) could not already involve a Pareto efficient choice of \( m(\sigma) \) and of side-transfers \( y_{m(\sigma)} \). Assuming a nonzero cost of renegotiation (even infinitesimal), there is then no reason for \( S \) to propose an \( EASC \) that she may want to renegotiate at the ex-post stage.

It remains to check that \( S \) will not want to wait for Stage 4 to offer the side-contract. Yet, any side-contract that is feasible at Stage 4 can be replicated by a side-contract designed at Stage 2, so there is no value in waiting. Thus, there is no loss of generality in considering only ex-ante side-contracts. ■

Indeed, \( N \) will always be misreported for \( G \) if \( E_{\sigma}^{e} t_{G} > E_{\sigma}^{e} t_{N} \) and \( E_{\sigma}^{e} s_{G} = E_{\sigma}^{e} s_{N} \). Then, \( P \) is strictly better off paying \( A \) the same amount when \( G \) turns up and when \( N \) does.
7.2 Supervisor’s Problem

Proceeding by backward induction we now solve for S’s problem. Suppose she offers an EASC inducing A into choosing \( e = \bar{e} \). Her problem is then of the form

\[
\max_{\{y_{\pi m(s)}\}} E_G^G (s_m(G) - K (y_m(G)) y_m(G)) + E_N^G (s_m(N) - K (y_m(N)) y_m(N)) \quad \text{s.t.} \\
E_G^G (t_m(G) + y_m(G)) + E_N^G (t_m(G) + y_m(G)) - \psi \geq u^A \quad (8) \\
E_G^G (t_m(G) + y_m(G)) + E_N^G (t_m(G) + y_m(G)) - \psi \geq E_N^G (t_m(G) + y_m(G)) . \quad (9)
\]

Constraint (8) guarantees that the agent participates in the EASC (and does not postpone to Stage 4, obtaining \( u^A \)). Constraint (9) ensures incentive compatibility of \( e = \bar{e} \).

We ignore limited liability constraints for brevity. It can be verified (ex-post) that the side contracts described, given the optimal GC presented below, do not violate any of them. Also, (9) is written anticipating that S will set \( y_{\pi m(B)} = -t_{\pi B} \) in order to punish A if deviating from the EASC. This is clearly optimal for S.

Suppose now S induces A into playing \( e = \bar{e} \). Her problem then is of the form

\[
\max_{\{y_{\pi m(s)}\}} E_B^G (s_m(B) - K (y_m(B)) y_m(B)) + E_N^G (s_m(N) - K (y_m(N)) y_m(N)) \quad \text{s.t.} \\
E_B^G (t_m(B) + y_m(B)) + E_N^G (t_m(B) + y_m(B)) \geq u^A \quad (10) \\
E_B^G (t_m(B) + y_m(B)) + E_N^G (t_m(B) + y_m(B)) \geq E_N^G (t_m(B) + y_m(B)) - \psi \quad (11)
\]

where (10) and (11) ensure, respectively, participation and incentive compatibility. (11) is written anticipating that S will set \( y_{\pi m(G)} = -t_{\pi G} \) in order to punish A when detecting a deviation.

S will set all the other side-transfers (i.e., those relating to the contingencies tak-
ing place out of equilibrium) to either zero or to sufficiently negative values to make unilateral deviations from her part unprofitable (as a commitment device).

Computing payoffs. Observe first that one may disregard Case 5. Indeed, setting \( E_a s_G > \max (E_a s_B , E_a s_N) \) leads to \( S \) systematically report \( G \). \( P \) can always do at least as well having \( E_a s_G = E_a s_N \) and \( E_a t_G = E_a t_N \), treating \( G \) and \( N \) as the same piece of information. We also ignore Case 4 as the locally optimal grand contract would be identical to that of Case 3, except for a strictly higher payment from \( P \) to \( S \) when the latter reports \( B \).

In all cases considered, it is optimal for \( S \) to set \( y_{\pi m(N)} = t_{\pi N} - t_{\pi m(N)} \). By so doing, \( S \) ensures that \( u^A \geq E_N (t_{m(N)} + y_{m(N)}) \) and \( u^A \geq E_N (t_{m(N)} + y_{m(N)}) - \psi \), in any of the Cases of Lemma 1. Therefore, \( \textit{[8]} \) is the relevant constraint when designing the EASC inducing \( \bar{e} \) and side-transfer \( y_{\pi m(G)} \) will be such that the constraint binds. Similarly, \( \textit{[10]} \) is always the relevant constraint when designing the EASC inducing \( \bar{e} \).

Side transfer \( y_{\pi m(B)} \) will be designed to bind such constraint.

Define now \( E^\bar{e}(s - y) \) as the expected payoff for \( S \) provided by the EASC inducing action \( e \in \{ \bar{e}, \underline{e} \} \). We have that

\[
E^\bar{e}(s - y) = E_G s_m(G) + E_N s_m(N) + K \cdot (E_G t_m(G) + E_N t_N - \psi - u^A) + K \cdot (E_N t_{m(N)} - t_N)
\]

\[
E^\underline{e}(s - y) = E_B s_m(B) + E_N s_m(N) + K \cdot (E_B t_m(B) + E_N t_N - u^A) + K \cdot (E_N t_{m(N)} - t_N)
\]

where we have omitted the argument of function \( K(.) \) to save on notation. Now that the optimal reaction of \( S \) has been computed for a given \( GC \), we turn our attention to \( P \)'s problem.

\[\textit{30}\] The reason is that, if side-contracting fails at Stage 2, \( A \) receives in each (equilibrium) state at least \( t_{\pi N} \), since \( S \)'s opportunistic behavior is tamed by the way \( P \) designs her salaries. Choosing \( y_{\pi m(N)} = t_{\pi N} - t_{\pi m(N)} \) thus allows her to maximise the rent extracted from \( A \) without compromising the latter's incentives to choose the intended action. One may suspect that a EASC inducing \( \bar{e} \) and such that \( y_{GG} = 0 \) and\( E_N y_{m(N)} = E_N (t_N - t_{m(N)}) \) - \( E_G t_G \) - \( E_N t_N \) + \( \psi \) + \( u^A \), that is plus what is needed to bind \( \textit{[8]} \), could be better from \( S \)'s perspective. However, this is actually never such that \( \textit{[8]} \) is binding, unless \( P \) commits to strictly higher expected payments than the ones we derive below. Since a better solution for \( P \) exists, this is never going to be the case.
7.3 Proof Lemmas 3, 4, 5 and 6

The following four lemmas allow us to simplify $P$’s problem, by providing restrictions on the set of implementable grand contracts.

**Lemma 3** $P$ always designs $GC$ such that

\[ E^e_{sB} \leq E^e_{sN} + \frac{1}{k} E^e_{sN} \leq E^e_{sN} + \frac{1}{k} E^e_{tG} \leq E^e_{sG} \leq E^e_{sN} + \frac{1}{k} E^e_{(t_G - t_N)} \]

**Proof.** We need to proceed considering each Case in turn. Suppose to be in Case 1. Given Assumption 1, the claim follows immediately. Suppose to be in Case 2. The expressions for $E^e(s - y)$ and $E^e(s - y)$ suggest the following. If $E^e_{sN} + \frac{1}{k} E^e_{tN} > \max (E^e_{sB}, E^e_{sG} + \frac{1}{k} E^e_{tG})$, then $m(\sigma) = N$ for any $\sigma$. The same outcome can be obtained by $P$ setting $E^e_{sN} = E^e_{sB}$ and $E^e_{tN} = E^e_{tB} = 0$ (treating $B$ and $N$ as the same piece of information). Expected payments when $\sigma = B$ would be strictly smaller than if $E^e_{sN} + \frac{1}{k} E^e_{tN} > E^e_{sB}$. If $E^e_{sB} > \max (E^e_{sN} + \frac{1}{k} E^e_{tN}, E^e_{sG} + \frac{1}{k} E^e_{tG})$, then $m(\sigma) = B$ for any $\sigma$. The same outcome can be obtained having $E^e_{sN} = E^e_{sB}$ and $E^e_{tN} = E^e_{tB} = 0$. Again, $P$ would pay strictly less. Therefore, the optimal $GC$ must be such that

\[ E^e_{sN} + \frac{1}{k} E^e_{tN} \leq \max \left( E^e_{sB}, E^e_{sG} + \frac{1}{k} E^e_{tG} \right) \]

and

\[ E^e_{sB} \leq \max \left( E^e_{sN} + \frac{1}{k} E^e_{tN}, E^e_{sG} + \frac{1}{k} E^e_{tG} \right) \]

The stated inequalities imply that either

\[ E^e_{sB} = \max \left( E^e_{sN} + \frac{1}{k} E^e_{tN}, E^e_{sG} + \frac{1}{k} E^e_{tG} \right) \]
or

\[
E^e_s G + \frac{1}{k} E^e t_G \geq \max \left( E^e_s N + \frac{1}{k} E^e t_N; E^e_s B \right)
\]

Recall now the two (alternative) conditions defining Case 2. Let us assume that \( E^e_s N > \max (E^e_s G, E^e_s B) \). If the first equality above holds, then, since it must be the case that \( E^e_s G + \frac{1}{k} E^e t_G \geq E^e_s N + \frac{1}{k} E^e t_N \) (otherwise \( E^e_s B = E^e_s N + \frac{1}{k} E^e t_N \), which cannot be true by assumption), it is easy to see that the claim is verified. If the second inequality holds, since \( E^e_s N + \frac{1}{k} E^e t_N > E^e_s B \) by assumption, it is also easily seen that the Lemma is verified. Let us assume now that \( E^e_s G = E^e_s N > E^e_s B \). The first equality above cannot be verified, since \( E^e_s G + \frac{1}{k} E^e t_G \geq E^e_s N + \frac{1}{k} E^e t_N \) but \( E^e_s B < E^e_s G \) by assumption. Then, since \( E^e_s N + \frac{1}{k} E^e t_N > E^e_s B \), the Lemma is verified. Consider now Case 3. If the first equality above holds, then conditions defining Case 3 imply that \( E^e_s G + \frac{1}{k} E^e t_G = E^e_s N + \frac{1}{k} E^e t_N = E^e_s B \). If the second inequality holds, it implies, combined with conditions defining Case 3, that \( E^e_s G + \frac{1}{k} E^e t_G \geq E^e_s N + \frac{1}{k} E^e t_N \geq E^e_s B \).

\[\Box\]

**Lemma 4** It must necessarily be the case that

\[
E^e(s - y) = \max_{m(N) = G, N} E^e_B s_G + E^e_N s_m(N) + \frac{1}{k} (E^e_B t_G + E^e_N t_m(N) - u^A).
\]

**Proof.** By Lemma 3, \( S \) prefers \( m(B) = G \) to \( m(B) = N \). Moreover, in Case 2, the sign of \( E^e_B t_m(B) + E^e_N t_N - u^A \) would be nonpositive, if \( m(B) = B \). This and Lemma 3 imply that \( S \) prefers \( m(B) = G \) to \( m(B) = B \). In Case 1 and Case 3, \( E^e_B t_B + E^e_N t_N - u^A \geq 0 \), so, using Lemma 3, \( S \)'s payoff is unchanged between \( m(B) = G, B \). Lemma 3 also implies that \( m(N) \neq B \). This implies that \( E^e_N t_m(N) \geq E^e_N t_m(N) - t_N \geq 0 \).

\[\Box\]

**Lemma 5** It must necessarily be the case that
$$E^e(s - y) = \max_{m(N) = G, N} E^e_G s_G + E^e_N s_m(N) + \frac{1}{k} \left( E^e_G t_G + E^e_N t_m(N) - \psi - u^A \right).$$

**Proof.** By Assumption 1 and Lemma 3, we have $m(G) = G$ and $m(N) \neq B$. This implies that $E^e_N t_m(N) \geq E^e_N (t_m(N) - t_N) \geq 0$. Consider now Case 1 and suppose $E^e_G t_G + E^e_N t_N - \psi - E^e_N t_N < 0$. Then constraint (12), using Lemma 4, can be written as

$$\max_{m(N) = G, N} E^e_G s_G + E^e_N s_m(N) + k \left( E^e_G t_G + E^e_N t_N - \psi - E^e_N t_N \right) + \frac{1}{k} E^e_N (t_m(N) - t_N) \geq \max_{m(N) = G, N} E^e_B s_G + E^e_N s_m(N) + \frac{1}{k} \left( E^e_B t_G + E^e_N t_m(N) - E^e_N t_N \right).$$

$P$'s problem is to minimize equilibrium payments subject to this constraint. Anticipating that $s_{GR} = t_{GR} = 0 \forall r$, then reducing $t_G$ and $t_N$ does not help to relax (12), since they appear with the same or higher frequency on the left than on the right hand side. Moreover, having $E^e_G t_G + E^e_N t_N - \psi - E^e_N t_N < 0$ can only tighten the constraint, as it would force $S$ to make a payment to $A$ in order to induce high effort, which has an extra cost of $k - 1$ per dollar transferred. Therefore, $P$ is strictly better off by having $E^e_G t_G + E^e_N t_N - \psi - E^e_N t_N \geq 0$. By a similar reasoning, it can be argued that $E^e_G t_G + E^e_N t_N - \psi - \sum_{\sigma} E^e_{\sigma} t_N \geq 0$ and $E^e_G t_G + E^e_N t_N - \psi - E^e_N t_N \geq 0$ must be optimal in Case 2 and Case 3 respectively.

**Lemma 6** Conditions $\sum_{\sigma} E^e_{\sigma} t_{\sigma} - \psi \geq E^e_N t_N$ (Case 1 and 3), $E^e_G t_G + E^e_N t_N - \psi \geq \sum_{\sigma} E^e_{\sigma} t_N$ (Case 2) are respected in any optimal GC.

**Proof.** Implicitly proven in Lemma 5.

### 7.4 Principal’s Problem

Making use of Lemma 3 to 5, the problem for $P$ can be written as
\[
\min_{\{s_G, t_G\}} \left\{ E_G (s_G + t_G) + E_N (s_{m(N)} + t_{m(N)}) \right\} \quad \text{s.t.} \quad \max_{m(N) = G, N} E_G s_G + E_N s_{m(N)} + \frac{1}{k} \left( E_G t_G + E_N t_{m(N)} - \psi - u^A \right) \geq (12)
\]

We ignore participation constraints for $S$ and $A$: given limited liability on transfers and normalizing the outside option for both to zero, they will always hold. Also, there is no need to write an incentive compatibility and a participation constraint for $A$, since they are necessarily respected as long as $P$ makes sure that (12) holds.

Constraint (12) is $S$’s incentive compatibility constraint. Transfers from $P$ have to be such that $E^e (s - y) \geq E^a (s - y)$. Using Lemma 3 and eliminating $u^A$ on both sides, (12) can be rewritten as

\[
E_G s_G + E_N s_G + \frac{1}{k} \left( E_G t_G + E_N t_G - \psi \right) \geq E_B s_G + E_N s_G + \frac{1}{k} \left( E_B t_G + E_N t_G \right)
\]

It is clearly optimal for $P$ to make payments to $S$ and $A$ only when high profits turn up, so we set $s_{\pi r} = t_{\pi r} = 0 \quad \forall r$. The problem then becomes, after simple rearrangements

\[
\min_{\{s_{\pi m}, t_{\pi m}\}} \left\{ P_G (s_{\pi G} + t_{\pi G}) + P_N (s_{\pi m(N)} + t_{\pi m(N)}) \right\} \quad \text{s.t.} \quad (2 \rho - 1) \left( s_{\pi G} + \frac{1}{k} t_{\pi G} \right) \geq \frac{\psi}{k}
\]

We will now propose a locally optimal $GC$ (solution to the above problem) for each of the 3 cases considered. Comparing $E(z)$ for each of them, we will determine the globally optimal one(s).

**Case 1.** Given the conditions defining Case 1, the solution to the problem is to set $s_{\pi G} = s_{\pi N} = s_{\pi B}$ and, since $N$ would otherwise be forged into $G$ by $S$, set $t_{\pi G} = t_{\pi N}$.

By Lemma 6, the solution summarizes as:
\[ t_{\pi G} = t_{\pi N} = \frac{\psi}{2\rho_\pi - 1 + (1 - \rho_\pi) p} > t_{\pi B} = 0, \]

\[ s_{\pi G} = s_{\pi N} = s_{\pi B} = \frac{\psi}{k} \cdot \frac{(1 - \rho_\pi) p}{(2\rho_\pi - 1)((2\rho_\pi - 1) + (1 - \rho_\pi) p)}, \]

with all other transfers to zero. This is GC “II” in the main text. It is such that

\[ E(z)^{II} = \rho_\pi \left( \frac{k-1}{k} \cdot \frac{\psi}{(2\rho_\pi - 1) + (1 - \rho_\pi) p} + \frac{\psi}{(2\rho_\pi - 1) k} \right). \]

**Case 2.** In this case \( P \) can differentiate salaries for \( S \) according to her reports. Optimality calls for \( t_{\pi N} = t_{\pi B} = 0 \), with \( s_{\pi N} = s_{\pi G} + \frac{1}{k} t_{\pi G} \) and, by the conditions defining Case 2, \( s_{\pi B} \in [0; s_{\pi N}) \). Note that \( P \) is indifferent as to the value of \( s_{\pi B} \) since \( B \) does not turn up on the equilibrium path. Using Lemma 6, the solution summarizes as

\[ t_{\pi G} = \frac{\psi}{pp_\pi} > t_{\pi B} = t_{\pi N} = 0 \]

\[ s_{\pi G} = \max \left( 0, \frac{1}{k} \left( \frac{1}{2\rho_\pi - 1} - \frac{1}{p_\pi} \right) \psi \right) \]

\[ s_{\pi N} = s_{\pi G} + \frac{1}{k} t_{\pi G} > s_{\pi B} \geq 0 \]

and all other transfers to zero. This is GC “III” in the main text. It is such that

\[ E(z)^{III} = \max \left( \left( \frac{1 - p}{p} \cdot \frac{\psi}{k} \right) : \left( \frac{\rho_\pi}{(2\rho_\pi - 1) - 1} \right) \frac{\psi}{k} \right) + \psi \]

**Case 3.** Using Lemma 6 and the conditions defining Case 3, one can see that the locally optimal GC would be the same as the one derived in Case 2, except that \( s_{\pi B} = s_{\pi N} \). Since \( B \) turns up only out of equilibrium, anyway, expected payments for \( P \) are the same as in GC “III”.

50
Global optimum. We are now in a position to select the optimal GC. Comparing GC II and III, one can verify that $E(z)^{II} > E(z)^{III}$ if $p \geq \frac{2\rho_π - 1}{\rho_π}$ or if $p < \frac{2\rho_π - 1}{\rho_π}$ and $k > \tilde{k}$, where

$$\tilde{k} \equiv 1 + \frac{2\rho_π - 1}{p(1 - \rho_π)} - \frac{p}{(1 - p)(2\rho_π - 1)}$$

Finally, if $p < \frac{2\rho_π - 1}{\rho_π}$ and $k \leq \tilde{k}$, then $E(z)^{II} \leq E(z)^{III}$.

Proof of Proposition 3

When $p \geq \frac{2\rho_π - 1}{\rho_π}$, “III” is the optimal GC when side-contracting is at both Stage 2 and 4, for any $k$. We have $E(z)^{III} < E(z)^I \iff k > k_1 = \frac{2\rho_π - 1 + (1 - \rho_π)p}{(1 - p)(2\rho_π - 1)}$.

When $p < \frac{2\rho_π - 1}{\rho_π}$, “III” is the optimal GC, when side-contracting is at both Stage 2 and 4, if and only if $k > \tilde{k}$. Otherwise, the optimal GC is “II”, which is always strictly more costly than “I”. We have $E(z)^{III} < E(z)^I \iff k > k_2 = \frac{2\rho_π - 1 + (1 - \rho_π)p}{p(1 - \rho_π)} > \tilde{k}$. Therefore, if $k < \tilde{k}$ we have $\max(E(z)^{II}, E(z)^{III}) = E(z)^{II} > E(z)^I$. If $k_2 > k \geq \tilde{k}$, we have $\max(E(z)^{II}, E(z)^{III}) = E(z)^{III} > E(z)^I$. Finally, if $k > k_2$, we have $\max(E(z)^{II}, E(z)^{III}) = E(z)^{III} < E(z)^I$.

Proof of proposition 4,5,6

To be completed.

References


Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift
für die gesamte Staatswissenschaft Vol. 146, No. 1, The New Institutional
Economics Different Approaches to the Economics of Institutions (Mar.
1990), pp. 85-105


Hierarchy. Journal of Law, Economics, & Organization # Vol. 6, No. 2
(Autumn, 1990) (pp. 301-324)

Annales d'économie et de Statistique. n.25/26, pp. 205-226


pages 280-305, Summer.

68(2), pages 309-342, March.


