Monitoring Team Production by Design*

Giovanni Bono
Università di Bologna – Dipartimento di Scienze Economiche
e–mail: bono@spbo.unibo.it
web: http://www.dse.unibo.it/bono/

June 10, 2005

Abstract

This work addresses the optimal design of the monitoring technology for a team when collective liability can not form. It shows that the principal's optimal design choice is then to concentrate monitoring on the less productive agent in a team. By controlling the less productive agent she fully discipline the more productive. This result helps in studying the interplay between the institutional set-up and the technological capabilities of teams.
Keywords: Team Production, Endogenous Information Structure, Individual Liability.
Jel: L23, C72, ...

 This document was produced on a GNU system (Debian GNU/Linux) using GNU Emacs (21.3) for the editing and LaTeX 2$_\varepsilon$ for the typesetting. Thanks to Debraj Ray. And to: Paolo Bertoletti, Alberto Bisin, Daniele Checchi, Inés Macho-Stadler, Pau Olivella, Howard Pethit. Last revision: June 28, 2004.

Contents

This work addresses the optimal design of the monitoring technology for a team when collective liability can not form. It shows that the principal's optimal design choice is then to concentrate monitoring on the less productive agent in a team. By controlling the less productive agent she fully discipline the more productive. This result helps in studying the interplay between the institutional set-up and the technological capabilities of teams.

This work studies the optimal design of team monitoring technologies. This is done using a team production model in the tradition of Alchian and Demsetz (Alchian and Demsetz, 1972:[**1**]) and Holmstrom (Holmström, 1982:[**3**]). That is a principal–multi agent game with hidden actions and production externalities. The standard framework is modified here in two respects. First, by adding a (non–degenerate) public monitoring technology. Agents actions are partially observable and the principal is in charge of monitoring design. Second, by changing the contracting technology available to the parties. Holmstrom's landmark contribution (Holmström, 1982:[**3**]) displays a sharing rule conditional on output that achieves efficiency. This holds in the standard framework regardless of other verifiable signals. Then a weakening of the contractual mechanisms is in order to study the optimal design of monitoring signals. To this end here the contracting stage game is defined so as to exclude the formation of collective liability. On one side, this rules out contracts that rely on one agent being accountable for another agent deviation—like the sharing rule in (Holmström, 1982:[**3**]). On the other, it points to a simple and general institutional feature.

Monitoring is modeled as a device that yelds informative signals only below certain treshold in the agents actions space. This reflects the inherent informational advantage of agents or a cost greater than the principal's gain above that treshold, and negligible below. The monitoring technology is a collection of alternative tresholds. The principal designs monitoring by choosing one of them. As an example, consider a device that can take recordings of one agent at a time. In this case, the principal chooses for how long to target each one agent facing the trade-off with the lenght of the other recordings.

In a two–agent model, the principal's optimal design choice is to concentrate monitoring on the less *productive* agent.[1] This is by definition the one whose least undetectable deviation from efficiency induces the lower loss in output. Controlling the less productive fully disciplines the more productive. The result rests on the possibility of combining observed output and cross-monitoring to infer contractible lower bounds on agents actions. The principal's best contracts are conditional on *direct* monitoring signals and on these *indirect* monitoring bounds. At least one agent is constrained by indirect more than by direct monitoring. Since indirect monitoring depends *ex post* on both agents actions, signing contracts puts agents in a game. For a given design, the agents strategic advantages are at the opposite ends because one's deviation would cause a lower output loss. In the Nash equilibrium of the game direct monitoring is binding for this agent while indirect monitoring is binding for the other. By choosing the design the pricipal essentially picks the agent who will deviate from efficiency.

The literature on principal many-agents problems explores mostly the stochastic properties of monitoring technologies. For example, in the cited seminal

[1] In a n-agents model the result holds with slight modifications. See the discussion in section 3.1.1.

contribution (Holmström, 1982:[**3**]) Holmstrom addresses the relation between these and incentives, concentrating on what makes information systems redundant. He shows that for a given set of observable signals the critical property is statistical sufficiency with respect to agent actions. This result is applied to the comparison of contracts that depend on an individual monitoring signal and contracts that depends on other agents signals as well, and the two are shown to be equivalent if agents outputs are independent. The characterization of classes of equivalence between (constrainedly) efficient and restricted mechanisms, in an analogous fashion, has been extensively pursued in the literature and is not reviewed here.

This work focuses instead on a restriction of the mechanisms available to the principal that precludes the formation of collective liability. With the exceptions of (Ishiguro and Itoh, 2001:[**5**]) and (Marx and Squintani, 2003:[**7**]) in the context of hidden information, the effects of impediments to collective liability is to my knowledge a neglected area of research. However, there are practical contracting environment in which its formation is not a trivial matter.[2] For example labor contracts provide tipically for individual liability, even when they are collectively bargained. Conversely, when they contain collective liability this is almost always partial and bargained in a heavily institutionalized environment.

With the proposed restriction in place it is possible to address the "physical" properties of monitoring as a simple allocation problem, leading to the corner solution which is the main claim of this work. This result agrees with a proposition advanced in an important contribution to the job–design literature (Hölmström and Milgrom, 1991:[**4**]). Holmstrom and Milgrom show in (Hölmström and Milgrom, 1991:[**4**]) that the optimal job design for two risk–averse agents implies the concentration of easy–to–monitor tasks on the one that bears the highest cost of effort. The asymmetry is induced by the strategic interaction between agents and regards the intensity of monitoring here, whereas is induced by their relative risk-preferences and regards its dispersion in (Hölmström and Milgrom, 1991:[**4**]).[3]

1 Framework

As in the standard team production model, agents—here denoted[4] $I = \{1, 2\}$—choose whether to sign a contract offered by the principal, who acts as a residual claimant and has all of the bargaining power. Both the principal and the agents utility indexes are linear in money, and production is represented by a smooth concave function. Agents actions—here denoted $\forall i : x_i \in [0, 1]$—are their private information unless in the scope of monitoring. Besides these, the information generated in production and monitoring is verifiable *ex post* by an enforcement authority trusted by the parties. Verification is based upon explicit

[2]See (Marx and Squintani, 2003:[**7**]) and for a discussion of these and other limits to collective liability

[3]Section 3.2.1 discusses the similarities and differences with their result.

[4]The notational conventions used here are: capital letters for sets, the corresponding small letters for elements, braces for unordered lists, and parentheses for ordered lists. An index preceded by "$-$" marks an element that is not the same as the one indexed without "$-$". An indexed element enclosed in parentheses stands for the complete list.

contracts.

1.1 Assumptions. The two assumptions presented hereafter depart from the standard model. The first specifies monitoring. The second introduces a a contracting stage not amenable to the formation of collective liability.

1.1.1 Monitoring. Monitoring delivers a *post*-production signal:

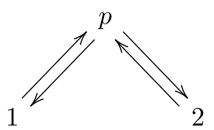
$$\forall i : \tilde{h}_i = \begin{matrix} x_i & \text{for } x_i < h_i \\ h_i & \text{for } x_i \geq h_i \end{matrix} \tag{1}$$

that fully reveal agent i's action below the treshold h_i and become uninformative above. The *design* of monitoring is a complete list of such tresholds: (h_i). The principal chooses the design from a given set $H \subseteq [0,1] \times [0,1]$. The requirement placed on H are that it contains a *component by component* upper bound:

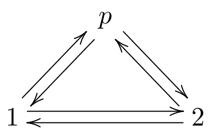
$$\forall i : \exists h_i^u (\forall h_i : h_i \leq h_i^u)$$

Note that H includes, as a degeneration, the standard case of unobservability of all agents actions.

1.1.2 Contracting Stage. Collective liability means that any agent bears responsability for one agent's deviation. To be held collectively liable, agents have thus to enter into two kind of commitments. One is a promise to act as guarantors for each other. The other is a promise to the principal. As a consequence of the first each agent is individually liable to each other. Collective liability is then build upon this web of mutual guarantees by the second one. Pictorially:



from this pattern of commitment collective liability can not form: each agent exchanges promises only with the principal ...



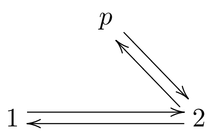
... here exchange of promises take place also between agents and collective liability may form.

In the standard team model the formation of collective liability is modeled as a rule of the stage game in wich agents sign contracts. At this stage, each agent has a veto power over production taking place: agents simultaneously sign a take it or leave it offer, and if one reject it the main game ends. The veto makes a contract conditional on any other contract being signed. That the principal is given all of the bargaining power means in this case that she chooses the agents mutual guarantees as well as their indiviual obligations to her. Without the veto acting as an automatic guarantee to others, collective liability can not emerge because of the multilateral threat of opportunistic behavior.[5] Then this

[5]Assuming that there is no alternative way of making contracts conditional on one another.

description of the contracting stage game excludes its formation: the principal choose whether to offer a contract to each agent and, if any is signed, production and monitoring take place. It is also assumed that a party who does not sign a contract get a reservation payoff normalized to zero.

Note the strategy to build collective liability through renegotiation adopted by (Ishiguro and Itoh, 2001:[**5**]). Ishiguro and Itoh assume in (Ishiguro and Itoh, 2001:[**5**]) that the principal uses only individual-based contracts. The timing of their game contains a further contracting stage in which one agent makes a take it or leave offer to the other and this becomes, if signed, their renegotiation proposal to the principal. See a more complete review of Ishiguro and Itoh's results in 3.2.2 and the picture of how collective liability takes form in their renegotiation stage here:



also from this pattern collective liability may form, but just one agent exchange promises with the principal as in the second contracting stage in (Ishiguro and Itoh, 2001:[**5**]).

1.2 Model. Output is in money units: $y = g((x_i))$ with $\forall i : \partial^2 g/\partial x_i^2 \leq 0 < \partial g/\partial x_i$. $g(1,1)$ is the efficien output. Denoting wages with w, utility indexes are: $v^p = y - \sum_i w_i$ and $\forall i : v_i = w_i - x_i$

1.2.1 Wage Rule. The principal offers a wage rule:

$$\forall i : w_i = \begin{matrix} \tilde{w}_i & \text{whenever } \tilde{h}_i = h_i \text{ and } y \geq g(1, \tilde{h}_{-i}) \\ 0 & \text{otherwise} \end{matrix} \tag{2}$$

This provides two criteria that a party can submit to an external authority for verification. If one fails, the rule defaults to no money transfer. The first criterion relies on monitoring directly. Signing this clause makes agent i commit to $x_i \geq h_i$. The second relies on cross-monitoring and on output. By signing this, agent i accepts responsability if output falls short of a value computed using $-i$'s monitoring signal. Since an agent would not commit to an unverifiable action on part of the other agent, that in (2) is the highest contractible level of output. Note that in the linear utility context of this work the contracts:

$$\forall i : w_i = \begin{matrix} \tilde{w}_i & \text{whenever } y \geq g(1,1) \\ 0 & \text{otherwise} \end{matrix}$$

induce the efficient actions if collective liabilty may form.

The second clause in (2) is computated inverting the production function at a point:

$$y \to \tilde{x}_i = g^{-1}(\cdot, x_{-1}) \text{ calculated in } x_{-1} = \tilde{h}_{-1}$$

and finding the level of output that maps to $\tilde{x}_i = 1$. The collective liabilty counterpart of this clause replaces the computation with:

$$y \to \tilde{x}_i = g^{-1}(\cdot, x_{-1}) \text{ calculated in } x_{-1} = 1$$

in which an agent acts as a guarantor if the other is doing the same.

Then both contracts need to include an analytical description of the production technology. This is a strong informational requirement, fulfilled in practice by contracts that link to blueprints, technical documentation, state of the art business practices, the organization past record and the like—or by enforcement authorities that use similar information in a predictable way. For the game presented here this requirement is of particular relevance, since it is not assumed that the monitoring design is revealed *ex ante* to agents.

As a benchmark, consider the exchange that would take place on a spot market. This is by definition limited to the verifiable (h_i). The efficiency gains above $g(h_i, h_{-i})$ are realized by contractual mechanisms thanks to the public character of technical knowledge in (2) or thanks to cross-guarantees when collective liability can emerge.

1.2.2 Agents Game. The optimal contracts $(\tilde{w}_i^*)$ are drafted by the principal taking into account the agents strategic interaction. It is convenient to describe the game in the optimal context, assuming that agents had the contracts signed and fulfill the obligations, and to describe how monitoring constrains the optimal actions of agents.

For a given (h_i), agent i respects the indirect clause of the contract if he chooses his optimal action in a region delimited by a reaction function:

$$x_i^* \geq \min\{x_i : g(x_i, x_{-i}^*) \geq g(1, h_{-i})\} \tag{3}$$

By construction, this coincides with the set above and including the isoquant of $g(\cdot)$ passing through $(1, h_{-i})$. At the same time, to meet the direct monitoring requirement:

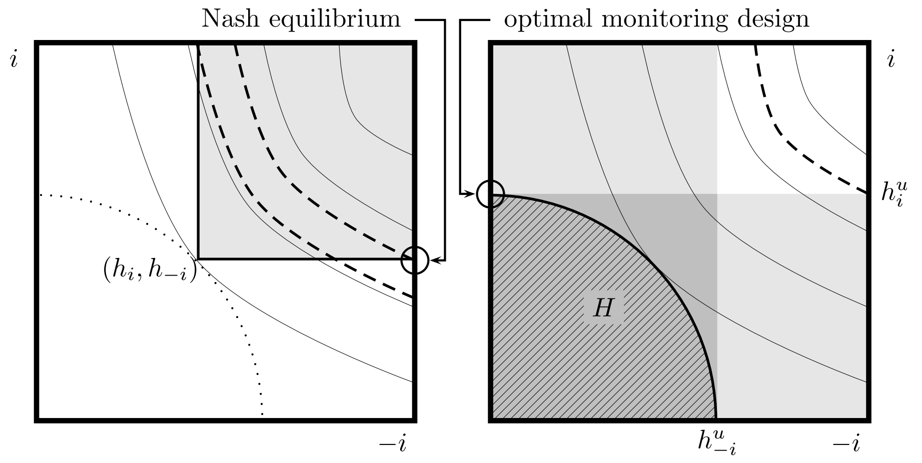
$$x_i^* \geq h_i \tag{4}$$

In the following the conditions in (3) and (4) are called the *indirect* and the *direct monitoring constraints*. Agent i chooses the minimal action that satisfy both and so $-i$. To find a Nash equilibrium of this game, assume that their indirect constraints are on different isoquants and hence can not cross. In this case, agent i's indirect constraint intersects $-i$'s direct constraint at just one point, where i's action reaches its maximal value. This point is an equilibrium if i's indirect constraint is on an higher isoquant than that of $-i$. Otherwise the equilibrium is the analogous point on $-i$'s indirect constraint. The Nash equilibrium of this game is thus unique unless the indirect monitoring constraints coincide. When this is the case, there is a continuum of equilibria along the isoquant.

In the unique equilibrium an agent is either on his direct or on his indirect monitoring constraint. Monitoring matters for just one agent and so the principal faces a choice between maximizing monitoring on one agent or on the other. The multiple equilibria case arises when she is indifferent between the two. Thus, the monitoring design for the unique equilibrium case is still optimal.

Figure 1 illustrates the strategic situation in two diagrams. Both plot a monitoring technology set H and some isoquants of a production function in the space of agents actions. In the example H is perfectly symmetric—$\forall a \in (0, 1) : (h_i = a) \in H \Leftrightarrow (h_{-i} = a) \in H$—and $g(\cdot)$ is such that $\forall a, b \in (0, 1) : a > b \Leftrightarrow g(a, b) < g(b, a)$.

Figure 1: the agents game and the problem of the principal



- ▷ the diagram on the left presents the agents game. Monitoring design is fixed at (h_i, h_{-i}). (3) and its analogous for agent $-i$ are the dashed isoquants. The lower belongs to agent i. (4) and analogous are both satisfied in the shaded region—including the frame. At the Nash equilibrium of the game agents actions are $(x_i^* = h_i, x_{-i}^* = 1)$. Note also that (h_i, h_{-i}) is on the highest isoquant crossing H. That is, is the design the principal would choose had she only access to a spot market.
- ▷ the diagram on the right presents the problem of the principal. The dashed isoquant corresponds to the optimal choice $(h_i^* = h_i^u, h_{-i}^* = 0)$. The shaded regions highlight the components upper bounds of H.

2 Optimal Monitoring Design

The principal solves the optimal design prolem looking for the higher $g(h_i^u, 1)$. Knowing that one agent is anyway playing $x_i^* = 1$, she maximizes monitoring on the other. Then, she maximizes monitoring on the agent that would cause the lower efficiency loss when he is on his direct monitoring constraint. In the following this potential loss is called *productivity*:

$$g_i^u = g(1, 1) - g(h_i^u, 1)$$

and a production function $g(\cdot)$ is called *anonymous* if $g_i^u = g_{-i}^u$. Anonymity depends on a single point in the domain of $g(\cdot)$ induced by the monitoring technology set upper-bounds. Thus for a given H an anonymous production function is non-generic.[6] The main claims of this work are then summarized:

[6] That is, is possible to find a function arbitrarily "close" to it but non-anonymous. For a work that uses genericity in the context of team production in a similiar altough more subtle way, and for the formal definitions, see (Battaglini, 2003:[**2**]).

Claim 1 (optimal monitoring design) There is an optimal design that contains the less productive agent's monitoring upper bound.

Claim 2 (unique optimal monitoring design) Unless the production technology is anonymous there is a unique optimal design.

These summarize also the best belief one agent can hold on the principal's monitoring design. Then is not relevant if the design choice is revealed *ex ante* or not. The proof[7] of the claims is given in four steps:[8]

Step 1 (claims) What is claimed must hold at the optimal contracts and for the optimal principal's and agent choices. Indicate with l the less productive agent—that is index I with $\{l, m\}$ so that: $g_l^u \leq g_m^u$. Then Claim 1 is:

$$(\tilde{w}_i^*) : (h_l^u, \cdot) \in \text{argmax}_{(h_i)}\ v^p(\forall i : x_i \in \text{argmax}_{x_i}\ v_i) \tag{5}$$

where the "$\cdot$" indicates a point in $\{h_{-i} : h_i^u \in H\}$ choosed according to some conventional rule. Claim 2 says that when $g_i^u \neq g_{-i}^u$ the first "$\in$" is replaced by "$=$" in (5).

Step 2 (agents game) Define the index set I_C so that an agent is in when his direct monitoring constraint (4) is binding at a Nash equilibrium: $\forall i : i \in I_C \Leftrightarrow x_i^* = h_i$. If an agent is in I_C he must be the less productive in a local sense:

$$\forall (h_i) : i \in I_C \Leftrightarrow g(h_i, 1) \geq g(1, h_{-i}) \tag{6}$$

($\Rightarrow$): if $x_i^* = h_i$ then by (3): $g(h_i, x_{-i}^*) \geq g(1, h_{-i})$; the conclusion follows because x_{-i}^* is bounded by 1. ✓

($\Leftarrow$): say $i \notin I_C$; then $x_i^* > h_i$ and again by (3) and from the premise: $g(x_i^*, x_{-i}^*) > g(1, h_{-i})$; since no constraint is binding x_i^* can not be in a Nash equilibrium, and the conclusion follows by contradiction. ↯

Note that from ($\Leftarrow$) also follows that I_C can not be empty.

Step 3 (Nash equilibrium)

$$(h_i, 1) \text{ is a Nash equilibrium} \Leftrightarrow g(h_i, 1) \geq g(1, h_{-i}) \tag{7}$$

($\Rightarrow$): say $g(h_i, 1) < g(1, h_{-i})$; then from (6) $-i \in I_C$. ↯

($\Leftarrow$): from (6): $i \in I_C$; then from (3): $x_{-i}^* = 1$. ✓

Step 4 (principal's problem) To prove Claim 1 say that $(h_l^u, \cdot)$ is not an optimal design for the principal. Using (7): $\exists i \in I : g(h_i, 1) > g(h_l^u, 1)$. ↯

To prove Claim 2 replace "an" with "the unique" and "$>$" with "$\geq$". ↯

Wages were never needed in the proof above. Once the partecipation constraints are met, agents optimal choices depends only on the monitoring design and

[7] A "✓" signals the conclusion of an argument, and a "↯" that the conclusion is reached by contradiction.

[8] Step 5 in section (3.1.1) extends the proof to the n-agents case.

the optimal monitoring design depends only on agents choices. To compute the optimal contracts substitute the Nash equilibrium that correspond to the principal's optimal monitoring design into (v_i), and fix agents utility at the reservation value. This yelds a constant optimal wage bill:

$$\sum_i \tilde{w}_i^* = 1 + h_l^u \tag{8}$$

3 Remarks

So, claim 1 and 2 establish the main result of this work:

$(\star)$ when collective liability can not form, the principal's optimal design choice is to concentrate monitoring on the less productive agent in a team.

To my knowledge, $(\star)$ is a new result in the literature on teams. Alternative specifications of the model's building blocks are discussed below (3.1), while 3.1.1 extends claim 1 and 2 to the n-agents case. Section 3.2 discusses the result on job design presented by Holmstrom and Milgrom in (Hölmström and Milgrom, 1991:[**4**]) and reviews Ishiguro and Itoh's renegotiation model (Ishiguro and Itoh, 2001:[**5**]). Section 3.3 concludes.

3.1 Model's Building Blocks.

- ▷ The monitoring technology H is a fixed set and the principal selects the design at no cost. One could take a collection of such sets instead, to shift focus on the cost structure of monitoring. See example 1.
- ▷ The monitoring signals described in (1) are perfectly informative below h_i. The substantive content of the assumption in section 1.1.1 is that they are uninformative above. Having monitoring signals of the form:

$$\forall i : \tilde{h}_i = \begin{matrix} \tilde{x}_i(x_i, x_{-i}) & \text{for } x_i < h_i \\ h_i & \text{for } x_i \geq h_i \end{matrix}$$

  where $\tilde{x}_i(\cdot)$ is a function, would make no difference for the equilibrium choices because of the direct monitoring constraints in (4).
- ▷ The agents utility indexes are linear in effort. Having them convex in effort[9] would only affect the value of (8).
- ▷ The environment is deterministic. In this context, the efficient outcome is reachable only in the trivial case in wich for all but one $i : h_i = 1$. Adding uncertainity in monitoring and/or in production could make $(\star)$ and full efficiency coexist if agent are sufficiently risk-averse. See example 2

Example 1 (Cost Structure) Consider the monitoring technologies $H_{(s_i)}$ of the form $\{(h_i) : h_i \leq s_i\}$. These are rectangular monitoring technologies, similar

[9] As usual in the literature on teams.

to the two–time shaded region in the right panel of figure 1. Since by (7) a Nash equilibrium of the agents game contains a component's upper bound, $H_{(s_i)}$ can be intended as a construct for the least costly technology among those characterized by the same component by component upper bound. Say that $H_{(s_i)}$ costs $c((s_i))$, with $c(\cdot)$ smooth and increasing in its arguments. The costs structure place a condition for $(\star)$ to hold on the entire collection $H_{(s_i)}$. Assuming $\forall a \in [0,1] : g(a,1) \leq g(1,a)$,[10] a simple way to characterize it is:

$$\forall a \in [0,1] : \min_{s_{-i}} c(a, s_{-i}) \leq \min_{s_i} c(s_i, a) + g(1,a) - g(a,1)$$

That is, at any point the minimal cost of monitoring the less productive agent must lie below the sum of the minimal cost of monitoring the more productive and of the difference in productivity between the two.

Example 2 (Efficiency) Let $(y, (\tilde{h}_i))$ be distributed according to a probability law over $((x_i), (h_i))$. Assume agents are equally risk-averse and their risk-aversion is constant in absolut terms. For a given $(\tilde{h}_i)$, call $F_y(\cdot)$ the cumulate density distribution over (x_i) induced by y. It is possible to write contracts based on $(\tilde{h}_i)$ and on a known relation:

$$(y, (\tilde{h}_i)) \rightarrow \tilde{x}_i = F_y(x_i | x_{-i} = \tilde{h}_{-i})$$

The obligation is then written $\max\{y : F_y(1|x_{-i} = \tilde{h}_{-i}) = 0\}$. By signing this agent i commit to the maximal level of output making *ex post* certain that $\overset{\text{as if}}{x_i} = 1$. Note that agents can not exchange guarantees and *a fortiori* they can not exchange insurance. If it was the case, for example, that the enforcement authority was adopting a known standard of evidence α, they could commit instead to $y : F_y(1|x_{-i} = \tilde{h}_{-i}) = \alpha$.[11] Then the principal offers contracts similar to (2) and the agents face constraints similar to (3) and (4). However, in computing the optimal actions the agents take into account the stochastic properties of the environment. These can be factored into the probability of being directly or indirectly constrained, and the distribution of observable signals. The optimal strategy for the indirectly constrained is again $x^*_{i \notin I_C} = 1$, and the principal offers him a wage that keep his partecipation constraint satisfied. If there is a probability of switching roles, she chooses (w^*_i) that make both agents behave as if they were the indirectly constrained and gets the efficient output. The optimal strategy for the directly constrained depends on the distribution of $\tilde{h}_{i \in I_C}$. If there is a positive probability that $\tilde{h}_{i \in I_C} = 1$, the principal can still find the incentive wage to induce the efficient behavior.

3.1.1 From Two to n-Agents. Claims 1 and 2 hold for any finite number of agents.

Step 5 (n-agents) Let I be $\{1, 2, \ldots n\}$ and modify $g(\cdot)$ and H accordingly. To extend the proof given in Steps 1 to 4 observe that:

$$\forall i : i \notin I_C \Leftrightarrow (-i) \in I_C$$

[10] Again as in figure 1.

[11] More generally the contractual probability should depend on the agents relative risk-aversion as well.

($\Rightarrow$): say $\exists x^*_{-i} > h_{-i}$; then: $g(x^*_i, (x^*_{-i})) > g(1, (h_{-i}))$. ↯

($\Leftarrow$): say $i \in I_C$; then: $g(x^*_i, (x^*_{-i})) < g(1, (h_{-i}))$. ↯

It is then possible to isolate two agents:

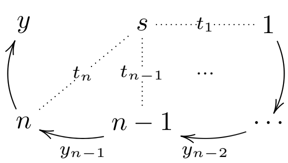
$$\begin{array}{lcl} \forall i : \mathsf{h^u_{-i}} & \leftarrow & (h_{-i}) \in \text{argmax}_{(h_{-i})}\{g(1, (h_{-i}))\} \\ \mathsf{m} & \leftarrow & i \in \text{argmax}_i\{g(1, \mathsf{h^u_{-i}})\} \\ \mathsf{l} & \leftarrow & i \in I/\{\mathsf{m}\} \end{array}$$

where $\mathsf{a} \leftarrow b$ means: pick one b and rename it a. After renaming[12] apply Step 1 to 4. ✓

As production involves the entire team, output can be combined with cross-monitoring to sort out just one contribution. The indirect monitoring constraint is binding for just one agent, and the principal maximizes direct monitoring on the remaining $n-1$. With more than two agents this portion of the optimal design—that is $\mathsf{h^u_{-i}}$—comes to depend on the shape of the production technology. In this sense adding agents makes the result loose strenght. At the same time, it envisions an incentive to slice many-agent production technologies into smaller processes. Note also that while with two agents only individual and collective liability are conceivable, with n agents $(2 + \ldots + n)$ different patterns of commitment may form.

Consider two examples of both ways of coping with the reduced scope of indirect monitoring in large teams. In both the more productive agent is a supervisor.

Example 3 (Assembly Line) For example of the first:



in an assembly line a supervisor (s) partecipate in n two–agent teams (t_i) with n line workers. Each team passes its output (y_i) to the next until the nth.

Here production is sliced sequentially.[13] The supervisor is the more productive agent in the large team and in any of the small teams. His contribution is the same whether he acts as one agent or he enters n different agency relations. Compare the end of the line output of the n small teams to that of a $(n+1)$-agents large team. The gain from slicing does not stem in this example from indirect monitoring *per se*. Rather, the gain is in the optimal monitoring design since:

$$g(1, (h^u_{-s})) \geq g(1, \mathsf{h^u_{-s}})$$

Example 4 (Hierarchy) An example of the second is a 3-agent team in which the supervisor takes responsability for $y \geq g(1, 1, h^u_l)$. All contracts contain disciplinary norms saying that an agent can be punished on the basis of monitoring *as if* the supervisor were exerting full effort—$x_s = 1$. If this condition is *for some reason* believed to be true, the gain stems in this example from:

$$g(1, 1, h^u_l) > g(1, \mathsf{h^u_{-s}})$$

[12] Note that there may be uncountably many such equivalent renaming.

[13] See (Strausz, 1999:[**9**]) for a model of sequential teams.

Here agents are grouped[14] and collective liability is formed using hierarchy: by signing contracts the other agents grant the supervisor disciplinary powers in exchange for his guarantees.

3.2 Related Literature.

3.2.1 Job Design.

As said ($\star$) bears a close resemblance to a result presented by Holmstrom and Milgrom This is in a section of their pioneering contribution to the literature on job design (Hölmström and Milgrom, 1991:[**4**]). They study the optimal grouping of tasks into jobs using a principal two–agent model. Risk-averse agents distribute attention t_i over a continuum of tasks k, incurring a cost convex in the total amount of attention $\int t_i(k)dk$. The attention they expend has a *per* task independent error variance that measures the task's "ease-of-monitoring". Any task can be shared by the agents, who are perfect substitutes for it. The principal can group tasks without restrictions and she offers to each agent a *per* task wage shedule. The main result—Proposition 5 in (Hölmström and Milgrom, 1991:[**4**])—state that when $\int t_1(k)dk < \int t_2(k)dk$ the optimal choice is to assign tasks "...so that all the harder-to-monitor tasks are undertaken by agent 1 and all the easiest-to monitor tasks are undertaken by agent 2".

Holmstrom and Milgrom proceed by showing that for a given total amount of attention "...it is never optimal for the two agents to be jointly responsible for any task k". Joint responsability makes the two agents bear the same risk. Since the optimal wage compensates for risk, splitting responsabilities lowers the wage bill. Then, they can use the convexity of agents costs in total attention to prove that is optimal to assign harder-to-monitor tasks to the agents that expend *less* total attention.

Both proposition 5 in (Hölmström and Milgrom, 1991:[**4**]) and ($\star$) are explanations for asymmetric distributions of monitoring. Note that Proposition 5 in (Hölmström and Milgrom, 1991:[**4**]) can live under the main assumption of this work because, as seen, there are verifiable monitoring signals and no complementarities in production.[15] Then these are not an issue in the comparison. Figure 2 highlights differences and similarities. The monitoring design is fixed in Proposition 5 in (Hölmström and Milgrom, 1991:[**4**]). By construction, the mean value of attention expended on a task is perfectly observed. Better monitoring is in this context less dispersed monitoring. Conversely, in ($\star$) the risk structure of monitoring plays no role. Better monitoring is here simply more monitoring. So the first result explains asymmetric distributions of the ease of monitoring, while the second explains asymmetric distributions of its intensity. The two distributions may well be orthogonal. The quantities that in equilibrium identify the better monitored agent are respectively total attention and productivity in $[H]$. These are different as:

- ▷ attention is an agent's cost measured in relative risk-aversion

[14] See (Matsushima, 2003:[**8**]) for a model of multi-group teams.

[15] That is, there is no scope for collective liability. Note that "joint responsability" for a task means in context that the task is split between agents and each agent is responsible for his portion.

Proposition 5 in (Hölmström and Milgrom, 1991:[**4**]):	($\star$):
for a given monitoring design over tasks	for a given job design
for a risk structure of tasks	for a production function
(with statistically independent tasks)	(continuous and twice differentiable)
which is independent of job design	which is independent of monitoring design
(with risk-averse agents and perfect substitutability between agents)	
the optimal job design …	the optimal monitoring design over jobs …
… **is to have better monitoring concentrated on** …	
… the agent with higher unit cost of effort	… the less productive agent

Figure 2: comparison with Holmstrom and Milgrom's result

▷ productivity is a principal's gain measured in money

The direction of the asymmetry is thus not commensurable. Assume for the sake of argument that harder-to-monitor jobs are also less intensely monitored. Then the direction of asymmetries is contradictory only if the agent that expends the highest amount of total attention is for some reason also the more productive in $[H]$. The highest amount of total attention is relative to two *ex ante* equally productive agents, and identifies the one that has the lower "unit cost of effort". If *ex ante* their productivity differs, the unit cost of effort of the more productive may as well be the lower. In the other direction, the more productive expends the efficient amount of effort in equilibrium. Since his wage is higher, his unit cost of effort is likely to be lower.

The most striking similarity between the two results is the role of "inherent nonconvexities" in causing the asymmetry. To both results can be applied Holmstrom and Milgrom's remark that "…since the ex ante specification of the model is symmetric in the roles of the two agents, if the problem entailed a concave objective and convex constraints, we would expect the optimal solution to be symmetric. However, the optimal solution … is not symmetric." The nonconvexities that sustain Proposition 5 in (Hölmström and Milgrom, 1991:[**4**]) are in the stochastic environment faced by the agents. Then job design is driven by efficient insurance against risk. Here the nonconvexities are in the ineherent uncertainity of the team actions faced by the principal. And, as seen before, monitoring design is driven by the agents strategic interaction.

Proposition 5 in (Hölmström and Milgrom, 1991:[**4**]) and ($\star$) are thus defined on different domains. One can however use them jointly to argue that under standard assumption on the agents risk-aversion the principal has an incentive to group easy-to-monitor tasks so that they form the less productive job in a team. This seems to fit the relation between productivity and monitoring-ease implied by the examples Holmstrom and Milgrom use to motivate their claim.

3.2.2 Renegotiation. Ishiguro and Itoh (Ishiguro and Itoh, 2001:[**5**]) study renegotiation in a principal two–agent model. After production, the agents observe a perfect but unverifiable signal that mutually reveals actions. The decentralized contracting stage described above in 1.1.2 follows this mutual monitoring signal. By assumption in the first contract stage the principal can not rely on unrestricted mechanisms. As shown for instance by Ma in (Ma, 1998:[**6**]), when agents actions are mutually observable there exists a mechanism that uniquely implement the first best action profile. Without the restriction in place, renegotiation would thus be pointless in (Ishiguro and Itoh, 2001:[**5**]) as much as monitoring is useless with collective liability contracts here. Ishiguro and Itoh's main result is that renegotiation "reduces the cost of implementing any implementable action profile down to the first-best level". This can be applied directly to $(\star)$ because they concentrate on initial constant-budget contracts as the one in (8).[16] The interest of the application is in that it envisions a route for modeling the endogenous formation of collective liability. As Ishiguro and Itoh argue, the mechanism decentrated *via* renegotiation is relevant for addressing unionized production environments.

3.3 Research directions. The result presented by this work:

$(\star)$ when collective liability can not form, the principal's optimal design choice is to concentrate monitoring on the less productive agent in a team.

is intended as an introductory step in the study of the relation between the institutional set-up and the technological capabilities of teams.[17]

The optimal choice is at a corner because the Nash equilibrium of the agents game—unique, except in a non-generic case—reduces the principal's problem to a binary choice. The skewness in the optimal design holds for an arbitrary monitoring technology set. The game between agents is a direct consequence of the set of identical contracts offered by the principal, in turn determined by the exclusion of collective liability. There is thus a direct link between the ability to form collective liability and the properties of the monitoring technology.

That this may be useful in the institutional study of the technological capabilities of teams is in a way obvious, as monitoring is itself a technological capability. More interestingly, monitoring collects and analyzes data on other technological capabilities. Since these data and analysis are relevant to contract enforcement one can expect them to be optimal, reliable and formal. Accordingly monitoring can be thought of as the output of an efficient research activity or as the ideal input for other research activities. In this vein, it generally lowers the cost of acquiring technological capabilities because it yelds research as a free by-product. Then the characteristics of the institutional set-up may induce different paths of technological learning in teams. $(\star)$ could imply, for example, specialization in the monitoring of low productivity jobs where the principal is not relying on collective liability and, by the external effect, the modularization of such jobs.[18] I am pursuing an examination of this line of interpretation and an attempt at

[16]When the first-best is not attainable with a constant-budget contract, they observe, it may "...exists an initial contract outside the set of constant-budget contracts that attain the first-best outcome *via* renegotiation". This is indeed the case here, if one adds the renegotiation stage and a variable component to the wage rule.

[17]It may as well be applied to regulatory issues and to other contexts.

[18]See example 3.

endogenizing the emergence of collective liability elsewher, as both issues require an extension of the present framework.

References

[1] A. A. Alchian and H. Demsetz. Production, information costs, and economic organization. *The American Economic Review*, 62:777–795, 1972.

[2] M. Battaglini. Moral hazard in teams with vector outputs. working paper, January 2003. `http://www.princeton.edu/~mbattagl/teams.pdf`.

[3] B. Holmström. Moral hazard in teams. *The Bell Journal of Economics*, 13(2):324–340, 1982.

[4] B. Hölmström and P. Milgrom. Multitask principal-agent analysis: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics & Organizations*, VII Sp:24–52, 1991.

[5] S. Ishiguro and H. Itoh. Moral hazard and renegotiation with multiple agents. *Review of Economic Studies*, 68:1–20, 2001.

[6] C. Ma. Unique implementation of incentive contracts with many agents. *The Review of Economic Studies*, 55(4):555–571, 1998.

[7] L. M. Marx and F. Squintani. Individual accountability in teams. working paper, University of Rochester, July 2003. `http://www.econ.rochester.edu/Squintani/teams.pdf`.

[8] H. Matsushima. Multi-group incentives. CIRJE Discussion Paper F-201, The University of Tokyo, March 2003. `http://www.e.u-tokyo.ac.jp/cirje/research/03research03dp.html`.

[9] R. Strausz. Moral hazard in sequential teams. *Journal of Economic Theory*, 85:140–156, December 1999.