Credibility and Monitoring: Outsourcing as a Commitment Device

Benjamin Bental[*], Bruno Deffains[†] and Dominique Demougin[‡§]

October 4, 2011

Abstract

We analyze an environment plagued by double moral hazard where the agent's effort level and the principal's precision in monitoring are not contractible. In such an environment, the principal tends to overmonitor thereby inducing low effort. To ease the latter problem, the principal may choose to increase monitoring costs by outsourcing the activity. As a result equilibrium monitoring is reduced and incentives become more powerful. This choice is particularly likely when the worker's effort is an important factor in determining output.

JEL H2

KEYWORDS Incentive scheme, integration, outsourcing, credibility

[*]Department of Economics, University of Haifa, Haifa, Israel and Department of Governance and Economics, EBS University, Wiesbaden, Germany, e-mail: bbental@econ.haifa.ac.il

[†]Ermes, University Paris 2 Panthéon Assas and Department of Governance and Economics, EBS University, Wiesbaden, Germany, e-mail: bruno.deffains@u-paris2.fr

[‡]Department of Governance and Economics, EBS University Wiesbaden, Germany, and University Paris 2 Panthéon Assas, e-mail: dominique.demougin@ebs.edu

[§]We thank an associate editor and two referees of JEMS for very useful comments. Demougin is grateful for hospitality and support of the University Nancy 2 and the BETA. Furthermore, financial support by the Deutsche Forschungsgemeinschaft under grant DE 1169/1-1 is gratefully acknowledged.

1 Introduction

Outsourcing in the private and public sectors seems to be on the rise (Holmström and Roberts 1998, Helpman 2006). Usually this trend is attributed to cost-gains. We propose a new characterization of this organizational choice, whereby a principal may choose to outsource an activity precisely because the costs associated with this choice are *higher*. Under such circumstances the outsourcing option may be used as a commitment device to increase credibility. Specifically, it increases monitoring costs and provides powerful incentives to govern the relationship between the principal and its agent.

Starting with Coase (1937) several key theories have been proposed to explain the boundaries of firms. The transactions cost theory associated in particular with Williamson focuses on the transaction as the basic explanatory unit (Williamson 1975, 1985). It suggests that transactions requiring complex contracts tend to be the ones that justify a non-market governance system. The property rights theory emphasizes the role of ownership of physical assets in an incomplete contracts environment (see Grossman and Hart 1986 and Hart and Moore 1990). Asset ownership provides bargaining power influencing negotiations in case of unforeseen events. Hence, the ownership structure influences expected returns and affects choices of investment into the relationship. According to this theory, the boundaries of the firm are drawn in a way that maximizes the parties' joint surplus. An alternative explanation of the boundaries of the firm hinges on the importance of asset ownership to align incentives in a multi-tasking environment (e.g. Holmström, 1999 and the literature therein). In such an environment, inefficiency arises because the agent's reward is not correctly aligned with the principal's objective. Giving the agent control over assets may mitigate the problem if the agent's ensuing incentives better coincide with those of the principal.

The current paper also studies the boundary decision of the firm as emerging from incentive considerations, albeit with a different mechanism based on double moral hazard. The agent's effort is not contractible. Hence, in order to induce effort the principal designs an incentive contract requiring the use of a proxy variable. The latter must be correlated with effort and is the outcome of monitoring. The precision of monitoring determines the strength of the correlation between effort and the proxy. Accordingly, the principal may trade off precision and the power of the contract in compliance with the agent's incentive compatibility constraint. Since in our setup precision is costly, the principal would like to reduce monitoring and increase power. However, precision is not contractible, resulting in double moral hazard. Specifically, contracts with more power will induce the principal to increase precision ex-post. Rational agents anticipate this implication and reduce effort.[1] We argue that under such circumstances the principal may find it useful to adopt an organizational form that increases monitoring costs, thereby increasing his own credibility.

We describe a setup in which integration is *a priori* informationally advantageous in monitoring. Nevertheless, we conclude that due to the aforementioned moral hazard on the part of the principal, the easier it is to obtain information within the organization, the more likely it becomes that the task will *not* be carried out within its boundaries. Intuitively, we are applying to the boundary of the firm a logic known since antiquity. Like Ulysses who anticipated his attraction to the sirens and asked to be tied to the mast, the principal in our environment anticipates his own future incentives to abuse monitoring and thus selects an organizational design that makes it costly for him to succumb to that temptation.

We assume that a decision to outsource an activity makes the collection of effort-related information more expensive by increasing both the level and marginal costs of monitoring. As a result, the principal faces a trade-off between organizational costs (implying different structures of monitoring costs) and the agent's induced performance. The extent of this trade-off depends on the impact of effort on the value created by the relationship relative to the organizational costs, leading to a counter-intuitive result: *ceteris paribus* the likelihood of outsourcing a task *increases* as the effort associated with that task becomes *more* valuable to the organization. In particular, this likelihood increases when the complementarity between the agent's effort and the principal's own productivity grows.

The adverse effect of the potentially opportunistic behavior of a principal lies at the center of the above argument. Such a possibility has long been recognized by the literature in conjunction with incentive contracts. For instance, one argument favoring tournament has been that it reduces gaming by the principal through fixing the sum of wages (see Malcomson 1984). A similar logic has been applied by Baiman and Rajan (1995) to contracts where an employer agrees to pay into a bonus pool and thereby eliminates any opportunistic action aimed at reducing bonus payments.[2] The recent financial crisis provides a number of real examples of such opportunistic behavior on the part of principals. For instance, in the dispute Commerzbank AG *vs.* Keen the employer tried to find *bona fide* reasons for not paying

[1] This setup provides a non-behavioral explanation to the negative effect of a principal's control on agents' performance experimentally found by Falk and Kosfeld (2006). We are grateful to a referee for pointing out this implication.

[2] Recently, Fisher *et al* (2005) carried out an experiment based on Baiman and Rajan's point where absent a precommitment on the total amout of boni to be allocated, "employees'" performace fell short of that achieved under precommitment.

a promised bonus (see Wynn-Evans 2007). The type of contractual incompleteness illustrated above forms the base of our analysis.

Opportunistic behavior in monitoring is analyzed also by Strausz (1997) who describes a double moral hazard problem where the principal cannot commit to a monitoring precision. Strausz suggests a different resolution based on delegation to a supervisor which results in the introduction of an additional contract. The latter provides a further instrument to resolve the double moral hazard. However, the possibility of collusion in the Principal-Supervisor-Agent game generates additional difficulties. Vafai (2005) finds conditions such that not delegating is preferable to the outcome of the Principal-Supervisor-Agent organization. Another possibility to deal with the principal's opportunistic behavior emerges in a repeated game framework. As is well known, this resolution works provided that the discount factor and the likelihood of the game breaking down remain sufficiently small.

Our paper contributes to several strands of existing literature. First, it is related to the Principal Agent literature which analyzes the implication of a standard moral hazard problem. Like the existing trade-offs between risk and efficiency (e.g. Holmström 1979), and rent and efficiency (e.g. Innes 1990 and Sappington 1983), we find that the principal induces too little effort compared to the first-best. Obviously, this is related to the double moral hazard literature (e.g. Al-Najjar 1997, Cooper and Ross 1985, Demski and Sappington 1991, Agrawal 2002). With respect to this literature, our paper is most closely related to Agrawal (2002) who also considers a tradeoff involving monitoring costs. Specifically, it provides conditions which generate one or the other type of incentive contracts. We extend the perspective of that paper by including sourcing issues. Second, our analysis belongs to the growing literature that uses microeconomic foundations to explain organizational choices (for a recent summary see Gibbons 2005). Third, our paper is related to the recent contributions on the association between sourcing decisions and firm productivity (e.g. Grossman and Helpman 2004, and Antràs and Helpman 2004). However, unlike our analysis, these papers which are based on the property-rights theory conclude that the *low* productivity firms tend to outsource, while the high productivity firms tend to integrate. Fourth, our analysis is also linked to the accounting literature on the double moral hazard problem. For example, Arya *et al* (1997) derive a set of conditions under which an accounting system that provides less public information may be desirable, as it raises the principal's commitment capability.

Our main argument is most closely associated with models that introduce credibility concerns to justify variations in the firm's boundary. In particular, Baker, Gibbons and Murphy (2002, 2005) impose a credibility requirement on implicit contracts in the context of a repeated game structure in order to derive the optimal organizational form. Even closer to our analysis, Kvaloy (2007) argues that outsourcing may be useful in preventing opportunistic behavior and enables the implementation of higher-powered incentives. In contrast to these papers, we analyze the impact of credibility on the design of explicit bonus contracts in a static framework without relying on repeated game arguments. Our analysis is also related to Ichino and Muehlheusser (2008) who consider the negative impact of too much monitoring in the context of an adverse selection problem. Here too an increase in the principal's marginal costs of monitoring has an advantageous effect.

Finally, our paper may be applied to the issue of outsourcing public services. The existing literature seems to form a consensus that hard-to-control and important tasks should *not* be outsourced (e.g. Hart, Shleifer and Vishny 1997, and Schmidt 1996). Similarly, Levin and Tadelis (2007) study outsourcing decisions of municipalities in the U.S. In line with most of the literature, they argue both theoretically and empirically, that harder to monitor services, and those whose quality is important, are more likely to be provided "in house". Our model shows that when opportunistic behavior on the part of the principal is likely, the second conclusion may be reversed.

The remainder of the paper is structured as follows. The next section presents the model. Section 3 contains the benchmark case with credible monitoring. Section 4 analyzes the double moral hazard case and derives implications for the organizational structure. Section 5 provides an example. Section 6 presents supporting evidence. Finally, section 7 offers some concluding remarks.

2 The Model

A principal owns a production technology that requires the input of an agent. Both parties are risk neutral and the agent's opportunity costs are zero. Moreover, the agent is not financially constrained. The value $v(a, \omega)$ of this relationship is increasing and concave in the agent's effort, a, and increasing in the technological productivity parameter ω that characterizes the principal. Moreover, effort and productivity are complements. Effort is costly with a monetary cost-equivalent given by $c(a)$, where $c'(a) > 0$, $c''(a) > 0$. Output and effort are not contractible, leading to a moral hazard problem on the part of the agent. Furthermore, while the results obtained by the monitoring technology used to align the agent's incentives are contractible, the principal's monitoring intensity is not; hence, generating a double moral hazard problem.

2.1 Monitoring and Contracting

Kim (1997) proved that in principal-agent setups with risk-neutral parties, verifiable information can be aggregated into a binary signal, $s \in \{0, 1\}$. Thus, in our framework the optimal incentive scheme will be a bonus contract without any loss of generality. Let θ denote the principal's monitoring intensity, and let

$$p(a, \theta) = \Pr[s = 1|a, \theta] \quad (1)$$

be the probability of observing the aggregated signal that induces the bonus payment.

Assumption: *The monitoring technology is such that* $p_a > 0$, $p_{aa} < 0$ *and* $p_\theta < 0$, $p_{\theta\theta} > 0$.

The requirement $p_\theta < 0$ makes sense in our context since we are interested in cases where the principal may (ab)use information opportunistically to reduce ex-post the likelihood of paying the bonus.[3] The additional convexity requirement means that this effect weakens as monitoring intensity increases.

To clarify, consider the following example. The likelihood of mistakes an agent may make in the course of his work decreases with effort. Mistakes can be detected by the principal depending on his monitoring intensity (for instance, the fraction of time spent on observing the agent's activity). In this case the bonus will be paid if the principal detects fewer mistakes than some critical level ($s = 1$). Holding the agent's effort constant while increasing the monitoring intensity necessarily raises the likelihood of detecting mistakes, thereby lowering the probability of paying the bonus.

The monitoring cost is a function $k(\theta, \gamma)$ where γ is a shift parameter with $k_\theta > 0$, $k_{\theta\theta} > 0$, $k_\gamma > 0$ and $k_{\theta\gamma} > 0$. The shift parameter γ stands for factors that increase both level and marginal monitoring costs, such as physical distance between the principal and the agent, complexity of the monitored tasks or organizational boundaries. In particular, in the sequel we associate integration ("I") and outsourcing ("O") with the parameters γ_I and γ_O, and assume that $\gamma_O > \gamma_I$.[4]

2.2 The timing of the game

Within the above environment, the parties make the following choices. The principal decides on the organizational form, designs a contract and chooses

[3] Demougin and Fluet (2001) give a set of requirements on the underlying information system such that $p_\theta < 0$. Intuitively, information systems should possess a non-crossing property in effort with respect to their informational content.

[4] For expository reasons, much of the analysis below is carried out taking γ to be a continuous variable.

monitoring intensity. The agent decides whether to participate, and if he does, how much effort to exert. The sequence of events is as follows. First, the principal decides whether to integrate or outsource. That choice then implies a corresponding monitoring technology and cost. Second, the principal makes a take-or-leave-it contract offer to the agent. Third, the agent accepts or rejects the offer. Fourth, assuming the agent has accepted the contract, the parties simultaneously make their respective choices, θ and a. Finally, the monitoring process generates the aggregate signal and payments are made according to the contract. The analysis follows the standard approach by applying backward induction.

3 Credible monitoring

Before turning to the analysis of the double-moral-hazard problem, we examine the simple benchmark case where the principal can credibly commit to a given monitoring intensity. In the above timing scheme, both the first and the fourth stages are affected. Specifically, in the first stage, the principal not only decides on the organizational form, i.e. γ, but also chooses the monitoring intensity θ. Consequently, in the fourth stage only the agent selects effort a.

The emerging setup generates a standard moral hazard problem with risk-neutral parties. At the contract design stage, the principal has already chosen γ and θ; hence, the monitoring intensity is already given. In such an environment the optimal contract specifies a fixed payment F and a bonus B to be paid solely when $s = 1$. Such contract generates an expected compensation $F + Bp(a, \theta)$ to an agent exerting effort a subject to a monitoring intensity θ.

Under these conditions, the principal can induce any effort level (including the first-best) by appropriately choosing B, and then extract the ensuing rent by adjusting F. Thus, the expected compensation associated with inducing effort a is simply $c(a)$. Accordingly, the principal's optimization problem at stage 1 becomes:

$$\max_{a,\theta,i\in\{I,O\}} \; v(a, \omega) - c(a) - k(\theta, \gamma_i). \tag{I}$$

Clearly, the principal will choose θ and γ so as to minimize the monitoring cost $k(\theta, \gamma)$. Given our assumptions above, monitoring intensity θ will therefore be as small as possible. In addition, the principal will select the organizational form that is associated with the lower monitoring and organization costs. As a consequence, with credible monitoring, the principal chooses integration, i.e. $\gamma = \gamma_I$.

4 The double-moral-hazard problem

We turn now to analyzing the original game, where the principal cannot commit to the choice of the monitoring intensity at stage 1. Here too we apply backward induction.

4.1 Effort and monitoring choice

Starting at the fourth stage where the agent has accepted the contract $\{F, B\}$ under the organizational structure γ, we determine the incentives of the two parties. At this stage, the relationship is characterized by a Nash game. Anticipating monitoring intensity, the agent chooses effort, while the principal chooses monitoring intensity expecting the agent's choice. Formally, the agent and the principal simultaneously solve:

$$\max_{a} F + Bp(a, \theta) - c(a) \tag{2}$$
$$\min_{\theta} F + Bp(a, \theta) + k(\theta, \gamma) \tag{3}$$

From the point of view of the principal's contract design problem, (2) represents the agent's incentive compatibility constraint, while (3) captures the principal's credibility requirement. The (2) constraint is concave. Therefore, the first order condition is sufficient. Analogously, the principal's objective (3) is convex, so that here too the first order condition is sufficient.

The resulting Nash equilibrium is characterized by the two resulting first order conditions:

$$Bp_a(a^N, \theta^N) - c_a\left(a^N\right) = 0 \tag{4}$$
$$Bp_\theta(a^N, \theta^N) + k_\theta(\theta^N, \gamma) = 0 \tag{5}$$

Implicitly, the first-order conditions yield effort and monitoring intensity as functions of the bonus, $a^N(B, \gamma)$ and $\theta^N(B, \gamma)$. Accordingly, this stage of the game generates a constraint on the "effort-monitoring" pair which the principal can implement. Notably, that pair depends on the organizational structure through γ.

These first-order conditions are equivalent to an $a - \theta$ contour condition $\mathcal{C}(a, \theta, \gamma) = 0$ where

$$\mathcal{C}(a, \theta, \gamma) \equiv \frac{p_a(a, \theta)}{p_\theta(a, \theta)} + \frac{c_a(a)}{k_\theta(\theta, \gamma)} \tag{6}$$

Assumption A1: $\mathcal{C}_a(a, \theta, \gamma) > 0$ and $\mathcal{C}_\theta(a, \theta, \gamma) < 0$ along the contour.

Geometrically, $A1$ implies that along the contour, monitoring intensity is implicitly an increasing function of the effort, hereafter $\theta^*(a,\gamma)$.[5] In the Appendix, we verify that $A1$ is based on a simple economic intuition with respect to the Nash equilibrium; it requires that the direct effects of a variation of the bonus in the respective agent's effort choice and the principal's monitoring decision are dominant. Specifically, consider the second order condition associated with (4). It implies that for a given θ, the level of effort is increasing in B. Similarly, (5) implies that when holding effort fixed, monitoring intensity increases in B. In equilibrium, these direct effects may be mitigated by feedback responses. In particular, with $p_{a\theta}(a,\theta) < 0$, increased monitoring intensity tends to reduce effort in the worker's decision problem. The conditions in $A1$ are equivalent to requiring that the feedback effects are dominated by the direct impact of changes in the bonus.

4.2 Contracting

At the third stage, the agent decides whether to accept the contract. Moreover, the agent correctly anticipates the outcome of the subsequent Nash game. Accordingly, given $a^N(B,\gamma)$ and $\theta^N(B,\gamma)$, the agent accepts the contract if his participation condition is satisfied.

In the second stage, the principal correctly anticipates both subsequent stages. Therefore, given any B and γ, the principal adjusts the fixed payment F to exactly meet the agent's participation constraint.[6] Taking this into account, the principal's objective simplifies to:

$$\pi(\gamma) = \max_B \quad v(a^N(B,\gamma),\omega) - c(a^N(B,\gamma)) - k(\theta^N(B,\gamma),\gamma). \tag{II}$$

The corresponding first-order condition is:

$$(v_a - c_a)\, a_B^N - k_\theta \theta_B^N = 0. \tag{7}$$

[5] One can observe here that the relation between effort and monitoring intensity (strategic complement or substitute) is complex. From the first order condition of the agent (equation (4)), we can solve for the optimal reaction of effort to changes in monitoring intensity. Similarly, from the first order condition of the principal (equation (5)), we can solve for the optimal reaction of monitoring intensity with respect to effort. While these go in opposite directions, the respective signs depend on the cross derivative of p with respect to these vatriables. This cross derivative can take any sign.

[6] Given that the fixed payment enters negatively the principal's profit, we know that F will be chosen to make the agent just indifferent between accepting and rejecting the contract. Substituting F into the principal's objective function yields the maximization problem (II).

From assumption $A1$, we know $a_B^N,\ \theta_B^N > 0$. Moreover, the marginal monitoring costs, k_θ, are also positive. Therefore, at the optimum we must have $v_a - c_a > 0$. This yields the following result:

Proposition 1 *The induced effort is too low, relative to the first-best.*

Notice that while under-investment in effort is standard in moral-hazard problems, the result here stems from a different mechanism. Specifically, in the existing literature inefficient effort may be due to one of the following trade-offs; risk vs. efficiency (e.g. Holmström 1979), rent vs. efficiency (e.g. Sappington 1983), incentives vs. aligning tasks efficiently (e.g. Holmström and Milgrom 1991), or equality vs. efficiency (e.g. Demougin, Fluet and Helm 2006). In our framework, none of these sources of inefficiency exists. Instead, under-investment follows because the principal cannot credibly commit not to over-monitor.

To see this, suppose the principal could pre-commit. Under these circumstances, the principal would find it optimal to reduce monitoring, thereby lowering the informational content of the signal. To align incentives, the principal would then offer a larger bonus. However, absent commitment, this is not credible. Therefore, in order to implement first-best effort, the principal would over-monitor. At the margin, it becomes useful to reduce effort, hence reducing the bonus as well as the principal's incentive to monitor excessively.

Our result can be related to findings in the behavioral literature on monitoring. For instance, in an experiment Falk and Kosfeld (2006) find that closer control of the agent hurts motivation while Ellingsen and Johannesson (2008) provide a theoretical framework consistent with the findings. Our model suggests a "non-behavioral" explanation of the negative relationship between monitoring and effort.

4.3 Organizational choice

The organizational choice made in the first stage clearly depends on the impact of γ on profit. In particular, if $\pi_\gamma < 0$, then outsourcing is a dominated option. On the other hand, when $\pi_\gamma > 0$, the principal may decide to outsource despite the higher organizational costs associated with this option. Given that a higher value of γ implies higher monitoring and organizational costs, one would expect the former case to prevail. We show, however, that the latter case may also occur and find conditions sufficient to generate $\pi_\gamma > 0$.

Applying the envelope theorem to (II), we find:

$$\pi_\gamma = (v_a - c_a)\, a_\gamma^N - k_\theta \theta_\gamma^N - k_\gamma \tag{8}$$

Using (7), we obtain:

$$\pi_\gamma = k_\theta \left[a_\gamma^N \frac{\theta_B^N}{a_B^N} - \theta_\gamma^N \right] - k_\gamma \tag{9}$$

The expression in the square brackets captures the total effect of a variation in γ on monitoring while taking the contour restriction into account. In Appendix 2, we show that under $A1$ this expression is positive. Thus, the entire first term on RHS of (9) is positive. It measures the indirect effect of a change in γ on profit. In contrast, the second term captures the direct impact of a change in γ on the principal's profit, which is clearly negative. In the same Appendix, we reformulate π_γ to derive the following result:

Proposition 2 *Outsourcing dominates integration whenever*

$$\frac{\varepsilon_{\gamma\theta}^k}{\varepsilon_{\theta\theta}^k + \varepsilon_{a\theta}^p - \varepsilon_{\theta\theta}^p} > 1, \tag{10}$$

where ε_{ij}^h denotes the elasticity of the function h with respect to the variables i and j.

Condition (10) considers a change in γ that has been normalized to generate a unit change in the effect of the direct cost. This is reflected by the RHS of the inequality. In contrast, the LHS measures the associated indirect effect. Accordingly, if condition (10) holds, the indirect effect dominates. In the following section, we provide an example to demonstrate that the above conditions may be met under circumstances that are not too strenuous.

Corollary 3 *The marginal effect of an increase in the principal's productivity, i.e. π_ω, is larger under outsourcing.*

Proof. Observe that $\pi_{\omega\gamma} = v_{\omega a}(a,\omega) a_\gamma^*$ where a_γ^* denotes the total effect of a variation in γ on effort. From the above argument, this effect is positive.[7] Thus, given the complementarity between the agent's effort and the principal's productivity, the result is obtained since $\gamma_O > \gamma_I$. ■

The corollary implies that outsourcing has a greater advantage for high productivity firms. As noted above, under outsourcing the higher credibility of the firm allows it to use more powerful contracts which leads to an increase in effort. Due to the complementarity, this effect is more beneficial for high productivity firms.

[7] The above argument is derived for the monitoring intensity θ^N. However, monitoring and effort are positively related along the credibility constraint.

4.4 Application to public services

Our setup can be applied to the provision of public services. In particular, it may help identify the circumstances under which a service should be publically provided or outsourced. For instance, consider a case where the high quality of a service is essential to avoid substantial damages (e.g. medical services, water purification, air-traffic control etc.). For these cases, the above setup needs to be slightly reinterpreted by setting $D(a,\omega) = -v(a,\omega)$ where D denotes the expected monetary value of damages and a the quality of the public service.

Suppose that producing quality is plagued by a moral hazard problem and that the public authority's ability to pre-commit to a given level of monitoring is limited. This may be due to ex-ante difficulties in contractually describing monitoring in full. In addition, elected public officials may have an incentive to over-monitor; for example, in order to temporarily reduce deficits or demonstrate their concern for public welfare.[8]

Given this structure, the public authority's objective is to minimize total costs under the correct incentive and credibility constraints. This implies that the requirement (6) must hold just as in the foregoing analysis. At the stage of deciding the organizational structure, the public authority's cost minimization problem becomes

$$\min_{a,\theta,\gamma} \; D(a,\omega) + c(a) + k(\theta,\gamma) \text{ s.t. } (6) \tag{III}$$

Clearly, problem (III) is equivalent to the above maximization setup. In general, the public authority's sourcing decision depends on the interpretation of ω. Below we provide a specific example concerning water provision in France and interpret ω as representing the importance of purification. We show that purification tends to be outsourced when its quality becomes crucial.

5 An example

In this section, we specify $p(a,\theta)$, $k(\theta,\gamma)$, and $c(a)$. The functional form describing the probability of obtaining the bonus builds on the example discussed in Section 2.1 and is directly taken from Demougin and Fluet (2001):

$$p(a,\theta) = a^{\theta}. \tag{11}$$

Next, we specify a quadratic cost function associated with monitoring

[8] Moreover, reputational considerations may be weak and influenced by the voting cycle.

and organizational structure:

$$k(\theta,\gamma) = \gamma\frac{\theta^2}{2} + f\theta + M(\gamma), \text{ with } f,\ M_\gamma > 0. \tag{12}$$

This functional form satisfies the initial requirements on $k(\theta,\gamma)$; It is increasing in θ and γ, convex is θ, and has a positive cross derivative. Furthermore, the function represents three types of costs. First, there is an interactive element that makes both the level and the marginal cost of monitoring increase as γ increases. Second, there is a monitoring cost element that is independent of γ. This captures the idea that some parts of the monitoring technology are independent of the organizational form. Finally, there is an organizational cost element that is independent of monitoring.

Turning to the effort-cost function for the agent, we specify:

$$c(a) = -\ln(1-a) \tag{13}$$

which is increasing and convex in a. Under the above specification the contour condition (6) yields:

$$\frac{\theta}{a\ln a} + \frac{(1-a)^{-1}}{\gamma\theta + f} = 0. \tag{14}$$

As is easily verified, $A1$ is satisfied. Solving for θ and keeping in mind that only the positive root is relevant, yields:

$$\theta = \frac{-f + (f^2 - 4\gamma a\,(1-a)^{-1}\ln a)^{1/2}}{2\gamma}. \tag{15}$$

Moreover, to obtain concavity of p with respect to a, monitoring intensity, θ, is restricted to be in the $[0,1]$ interval. For that purpose and in order to guarantee an interior solution, we choose the underlying parameters appropriately.

With $M(\gamma) \equiv 0$, we verify in Appendix 3 that the above specification satisfies the conditions underlying proposition 2, making outsourcing the dominating organizational form. However, reintroducing the organization cost $M(\gamma)$ may overturn this conclusion. Whether it does or not depends predominantly on the principal's productivity parameter ω and its interaction with the agent's effort. Specifically, when ω and a are complements, an increase in the principal's productivity parameter implies that the agent's contribution to profit becomes more important. Consequently, one would expect outsourcing to be more attractive since it raises credibility. In order to demonstrate these effects below, we specify the value generated by the match between the principal and the agent to be

$$v(a,\omega) = 2\omega a^{1/2} \tag{16}$$

5.1 Numerical specification

In equation (12) we set $f = 1$, and examine two values for the γ parameter; $\gamma = 1$ for the organizational structure with integration and $\gamma = 4$ associated with outsourcing. For the geometric representation, we let the productivity factor ω range between about 0.4 and 1.5. In line with the results stated above, the effort-choice curves are increasing in the productivity parameter ω. Moreover, increasing the marginal cost of monitoring by raising γ causes the principal to induce higher effort at every productivity level which reflects the weakened moral hazard on the part of the principal. Monitoring also increases in the productivity parameter, but a higher γ reduces monitoring intensity.

Internet Files/Content.Outlook/H4GTJR4K/graphics/fig01.gif

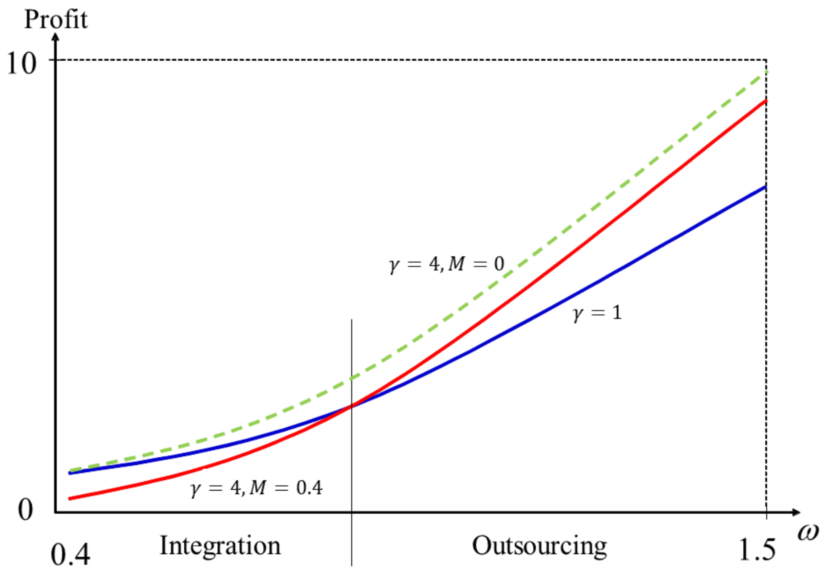

Figure 1: Productivity Costs and Sourcing

Figure 1 shows the positive impact of changes in ω on profits. Note that in the current example, the condition of proposition 2 holds. Hence, the indirect effect of γ on profit dominates its direct effect on cost. Specifically, with $M(\gamma) \equiv 0$, profits are uniformly higher under outsourcing for all ω and the principal always prefers to outsource the activity. Observe in addition that in accordance with the corollary, the slope of the profit curve with respect to ω is larger in the case of outsourcing. For this very reason the introduction of organizational costs (in Figure 1, $M(4) = 0.4$) causes the resulting profit curve to cross the profit curve associated with integration

($\gamma = 1$) from below, at about $\omega = 0.7$. Therefore, for all values of ω that are smaller than 0.7 the principal chooses to integrate, whereas for larger values of the productivity parameter, outsourcing becomes the dominant strategy.

6 Some Supporting Evidence

The following examples describe organizational choices that are hard to reconcile with most of the existing literature on the sourcing decision. While we do not claim that these observations provide conclusive "proof" corroborating our model, we do argue that our model's conclusions are consistent with the evidence we present.

6.1 Japanese car makers

In their review paper, Holmström and Roberts (1998) provide numerous cases of business relationships with disintegrated organizational structures that appear puzzling from the point of view of existing theories. For example, they point out that unlike American car makers, Toyota outsources design-intensive tasks. Since the organizational form of the American car makers fits well with standard predictions, Toyota's decision to deviate from it requires explanation. Holmström and Roberts suggest that in the Japanese system "... the automakers are inherently too powerful and thus face too great a temptation to misbehave opportunistically". In that respect, they argue that outsourcing "raises the cost of misbehavior" and thus increases credibility.

The model presented above follows a similar logic; outsourcing raises the marginal monitoring costs and reduces the incentives of opportunistic behavior in monitoring which increases the firm's credibility. Moreover, the firm will be willing to pay the high cost of outsourcing precisely when the task involved is important.

6.2 Regulation

This example applies the same idea to the sphere of government regulations. Gilardi (2005) claims that there is an increasing trend of governments to outsource regulatory activities in broad spheres of economic undertakings; e.g. utilities, financial markets, pharmaceuticals, food safety and environment. Girardi argues that this trend may be due to "the need for governments to increase their credible commitment capacity" which "lead(s) them to delegate regulation to an agency that is partly beyond their direct control."

Here again, the intuition is related to our analysis. Using the language of our paper, by delegating regulation to an agency, a government raises its cost of opportunistic behavior. In order to overturn a regulatory action, the government would need to work through the independent regulatory agency or, in an extreme case, abolish it. These are politically costly activities which weaken the incentive to renege.

6.3 Water provision in France

This example considers the sourcing decision of a specific service, namely of water purification in France, and enables us to match the model's structure with an actual sourcing problem. According to French law, municipalities are responsible for supplying drinking water as well as the collection and treatment of sewage. They are allowed to either delegate water management services to a specialized private company or keep it under their direct control. However, independently of this sourcing decision, municipalities remain responsible for water quality.

The final quality of drinking water depends on the initial condition of the source water and the treatment process. For instance, underground water is generally pure, requiring little treatment, whereas surface water tends to be of poorer quality necessitating complex treatment and meticulous control. Different municipalities face different types of water sources. Naturally, these variations affect the trade-offs embedded in the principal-agent relationship. Usual transaction cost arguments based on the complexity of contracts or the importance of the task involved suggest that municipalities should outsource simple cases while integrating the complex ones. In contrast, as stated in 4.4 our model predicts that as the importance of purification increases, outsourcing should become more likely.

The data we have contain 3606 French municipalities for which we know whether water supply is operated by the municipality or outsourced.[9] In addition, we know for each of these municipalities to what extent the water undergoes treatment. There are five levels of treatment, ranging from no disinfection to heavy disinfection. We also know whether water originates underground.

Figure 2 summarizes the data. As can be seen, a systematic bias in the distribution of water quality according to the mode of operation of the water

[9] The data was collected by the INSEE. There are 29,000 contracts from which the ones analyzed have been selected by the INSEE. The selection contains all municipalities with more than 10,000 inhabitants and a random sample of the smaller municipalities. While the data contain information on the type of the contract involved, unfortunately no further details on the incentive structure are available.

systems seems apparent. Whereas the high quality water source seems to be more frequent among the publicly owned operators, the medium-light and medium-heavy disinfection procedures are clearly more frequent among the privately owned operators than in the population, and therefore these categories are "under represented" among the publicly owned operators. This is also true for easily treatable water, albeit less clearly. The same holds for the underground water source, which is relatively easy to treat. The only exception to this pattern occurs for the worst water quality, where the public sector seems to have slightly more than the population share. However, it should be noted that the latter difference is statistically insignificant.[10]

Internet Files/Content.Outlook/H4GTJR4K/graphics/Fig1.wmf

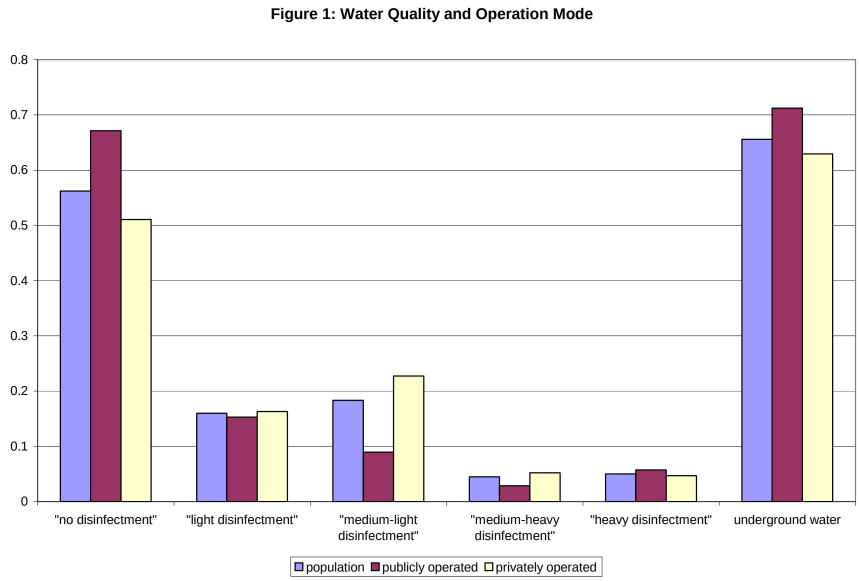

Figure 2: Data Characteristics

Altogether the empirical evidence in this case is consistent with the predictions of our model. The tendency to outsource water purification *increases* as the source becomes more contaminated. According to our model, this makes purification more important (in terms of the responsibilities of the public authority) which increases the need to obtain high-quality results. Outsourcing enables municipalities to use high-powered contracts and obtain the desired effect.

[10]We have checked the impact of population size on the sourcing decision and found it to be highly insignificant.

7 Conclusion

This paper shows that when a principal cannot commit to a monitoring policy, he may benefit from increasing monitoring costs through outsourcing. The paper argues that this conclusion is in line with some outsourcing decisions of firms which otherwise appear puzzling. Specifically, one would expect that higher monitoring costs associated with outsourcing would increase the tendency to integrate. We show that this argument against outsourcing can be turned on its head in the presence of double moral hazard. Deriving conditions under which this occurs, the paper provides a novel trade-off to answer Coase's (1937) original question as to which transactions are more efficiently conducted within a firm rather than through a market relationship. In particular, we conclude that outsourcing dominates when the agent's effort is sufficiently beneficial to the principal.

We have presented some evidence related to the sourcing decisions. The sourcing patterns of French municipalities concerning water treatment appear not to be in line with predictions found in the existing literature, but are consistent with our theoretical framework. Furthermore, our conclusions may also help explain other observed organizational choices that are puzzling from the point of view of most alternative theories discussed by Holmström and Roberts (1998). Their resolution of the contradictions is based on repeated games arguments, while we suggest that credibility issues may generate the observed phenomena even in a static framework. Finally, unlike other analyses, the model generates a positive association between high-productivity and the tendency to outsource. This positive association also seems to be consistent with data (see, e.g. Olsen's (2006) OECD survey).

Bibliography

Agrawal, Pradeep (2002): "Double Moral Hazard, Monitoring, and the Nature of Contracts," *Journal of Economics*, 75(1), 33-61.

Al-Najjar, Nabil (1997): "Incentive contracts in two-sided moral hazards with multiple agents," *Journal of Economic Theory*, 74(1), 174-195.

Antràs, Pol and Elhanan Helpman (2004), "Global Sourcing," *Journal of Political Economy* 112(3), 552-580.

Arya, Anil, Jonathan Glover and K. Sivaramakrishnan (1997): "Commitment Issues in Budgeting", *Journal of Accounting Research*, 35, 273–278.

Baiman, Stanley and Madhav V. Rajan (1995): "The Informational Advantages of Discretionary Bonus Schemes", *The Accounting Review*, 70(4), 557-579.

Baker, George, Robert Gibbons, and Kevin J. Murphy (2002): "Relational Contracts and the Theory of the Firm", *Quarterly Journal of Economics,* 117, 39-83.

Baker, George, Robert Gibbons, and Kevin J. Murphy (2005): "Strategic Alliances: Bridges Between 'Islands of Conscious Power'" Unpublished manuscript, USC's Marshall School.

Coase, Ronald (1937): "The Nature of the Firm", *Economica*, 4, 386–405.

Cooper, Russell and Thomas Ross (1985):"Product Warranties and Double Moral Hazard", *RAND Journal of Economics*, 16(1), 103-113.

Demougin, Dominique and Claude Fluet (2001): "Monitoring versus Incentives", *European Economic Review*, 45(9), 1741-1764.

Demougin, Dominique, Claude Fluet and Carsten, Helm (2006): "Output and Wages with Inequality Averse Agents", *Canadian Journal of Economics*, 39(2), 399-413.

Demski, Joel and David Sappington (1991): "Resolving Double Moral Hazard Problems with Buyout Agreements", *RAND Journal of Economics*, 22(2), 232-240.

Ellingsen, Tore and Magnus Johannesson (2008): "Pride and Prejudice: The Human Side of Incentive Theory", *American Economic Review*, 98(5), 990-1008.

Falk, Armin and Michael Kosfeld (2006): "The Hidden Costs of Control", *American Economic Review*, 96(5), 1611-30.

Fisher, Joseph G., Laureen A. Maines, Sean A. Peffer and Geoffrey B. Sprinkle (2005): "An Experimental Investigation of Employer Discretion in Employee Performance Evaluation and Compensation", *The Accounting Review*, 80(2), 563-583.

Gibbons, Robert (2005): “Four Formal(izable) Theories of the Firm?”, *Journal of Economic Behavior & Organization*, 58, 200-245.

Gilardi, Fabrizio (2005): “The Formal Independence of Regulators: A Comparison of 17 Countries and 7 Sectors”, *Swiss Political Science Review*, 11(4), 139-167.

Grossman, Sanford, and Oliver Hart (1986): “The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration,” *Journal of Political Economy*, 94, 691–719.

Grossman, Gene and Elhanan Helpman (2004): "Managerial incentives and the international organization of production", *Journal of International Economics,* 63, 237– 262.

Hart, Oliver, and John Moore (1990): “Property Rights and the Nature of the Firm,” *Journal of Political Economy*, 98, 1119–1158.

Hart, Oliver, Andrei Shleifer and Robert W. Vishny (1997) , „The Proper Scope of Government: Theory and an Application to Prisons,“ *Quarterly Journal of Economics*, 112(4), 1127-1161.

Helpman, Elhanan (2006): "Trade, FDI, and the Organization of Firms", *Journal of Economic Literature*, 44(3), 589-630.

Holmström, Bengt (1979): "Moral Hazard and Observability", *Bell Journal of Economics*, 10, 74-91.

Holmström, Bengt (1999): "The Firm as a Subeconomy", *Journal of Law, Economics, and Organization*, 15, 74-102.

Holmström, Bengt and Paul Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, & Organization*, 7, 24-52.

Holmström, Bengt and John Roberts (1998): "The Boundaries of the Firm Revisited", *Journal of Economic Perspectives*, 12(4), 73-94.

Ichino, Andrea and Gerd Muehlheuser (2008): "How often should you open the door? Optimal monitoring to screen heterogeneous agents", *Journal of Economic Behavior & Organization*, 67(3-4), 820-831.

Innes, Robbie (1990): "Limited Liability and Incentive Contracting with ex ante Choices", *Journal of Economic Theory*, 52, 45-67.

Kvaloy, Ola (2007): "Asset Specificity and Vertical Integration", *Scandinavian Journal of Economics*, 109(3), 551-572.

Kim, Son Ku (1997): "Limited liability and bonus contracts", *Journal of Economics and Management Strategy*, 6, 899-913.

Malcomson, James M, (1984): "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy* 92, 486-507.

Levin, Jonathan and Steven Tadelis (2007): "Contracting for Government Services: Theory and Evidence from U.S. Cities," NBER Working Papers 13350.

Olsen, Karsten Bjerring (2006): "Productivity Impacts of Offshoring and Outsourcing: A Review", Organization for Economic Co-Operation and Development, STI Working Paper 2006/1.

Sappington, David (1983): "Limited Liability Contracts between Principal and Agent"; *Journal of Economic Theory*, 29, 1-21.

Schmidt, Klaus (1996): "The Costs and Benefits of Privatization: An Incomplete Contracts Approach", *Journal of Law, Economics, & Organization*, 12(1), 1-24.

Strausz, Roland (1997): "Delegation of Monitoring in a Principal-Agent Relationship", *The Review of Economic Studies*, 64, 337-357.

Vafai, Kouroche (2005): "Collusion and Organization Design,", *Economica*, 72, 17-37.

Williamson, Oliver (1975): *Markets and Hierarchies: Analysis and Antitrust Implications*, New York, The Free Press.

Williamson, Oliver (1985): *The Economic Institutions of Capitalism*, New York, The Free Press.

Wynn-Evans, Charles (2007): "Discretionary Bonus Awards, UCTA and the Duty to Give Reasons", *Industrial Law Journal,*36, 207-213.

Appendix 1

Claim 4 *A1 is equivalent to $a_B, \theta_B > 0$.*

Proof. Taking total derivative of the first-order conditions (4) and (5), we obtain:

$$\begin{pmatrix} a_B \\ \theta_B \end{pmatrix} = \frac{-1}{\det} \begin{pmatrix} Bp_{\theta\theta} + k_{\theta\theta} & -Bp_{a\theta} \\ -Bp_{a\theta} & Bp_{aa} - c_{aa} \end{pmatrix} \cdot \begin{pmatrix} p_a \\ p_\theta \end{pmatrix}$$

Given that the determinant, det, is negative, we have:

$$\operatorname{sign}(a_B) = \operatorname{sign}\left[\left(Bp_{\theta\theta} + k_{\theta\theta}\right) p_a - Bp_{a\theta} p_\theta\right]. \tag{A1}$$

Suppose $a_B > 0$, then

$$Bp_{\theta\theta} p_a + k_{\theta\theta} p_a - Bp_{a\theta} p_\theta > 0, \tag{A2}$$

and substituting B from the principal's first-order condition, (A2) becomes:

$$-\frac{k_\theta}{p_\theta} p_{\theta\theta} p_a + k_{\theta\theta} p_a + k_\theta p_{a\theta} > 0. \tag{A3}$$

Next, substituting p_a from the $a-\theta$ contour condition and multiplying the result by p_θ / k_θ (which is negative), we obtain:

$$\frac{p_{a\theta}}{p_\theta} - \frac{p_{\theta\theta} p_a}{(p_\theta)^2} - k_{\theta\theta} \frac{c_a}{(k_\theta)^2} < 0 \tag{A4}$$

Finally, note that the LHS of (A4) is $\mathcal{C}_\theta$.

The equivalence between $\mathcal{C}_a > 0$ and $\theta_B > 0$ follows a similar argument. ■

Appendix 2

Claim 5 $a_\gamma^N \frac{\theta_B^N}{a_B^N} - \theta_\gamma^N > 0$

Proof. Consider the $a-\theta$ contour along the Nash equilibrium path $\mathcal{C}(a^N(B,\gamma), \theta^N(B,\gamma), \gamma) \equiv 0$. Taking total derivatives with respect to B and γ yields

$$\mathcal{C}_a a_B^N + \mathcal{C}_\theta \theta_B^N = 0 \tag{A5}$$

$$\mathcal{C}_a a_\gamma^N + \mathcal{C}_\theta \theta_\gamma^N + \mathcal{C}_\gamma = 0 \tag{A6}$$

Using these equalities, we find:

$$a_\gamma^N \frac{\theta_B^N}{a_B^N} - \theta_\gamma^N = \frac{\mathcal{C}_\gamma}{\mathcal{C}_\theta} > 0 \tag{A7}$$

■

Claim 6 $\pi_\gamma > 0$ *iff*

$$\frac{\varepsilon_{\gamma\theta}^k}{\varepsilon_{\theta\theta}^k + \varepsilon_{a\theta}^p - \varepsilon_{\theta\theta}^p} > 1.$$

Proof. Taking total derivatives of (4) and (5) with respect to γ yields

$$\begin{pmatrix} a_\gamma^N \\ \theta_\gamma^N \end{pmatrix} = \frac{-1}{\det} \begin{pmatrix} Bp_{\theta\theta} + k_{\theta\theta} & -Bp_{a\theta} \\ -Bp_{a\theta} & Bp_{aa} - c_{aa} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ k_{\theta\gamma} \end{pmatrix}$$

Using the implied results together a_B and θ_B from Appendix 1 in (9), we find

$$\pi_\gamma = k_\theta \left[\frac{k_{\theta\gamma} p_a}{(Bp_{\theta\theta} + k_{\theta\theta}) p_a - (Bp_{a\theta}) p_\theta} \right] - k_\gamma \tag{A8}$$

Using (5) to eliminate B and dividing by $k_\theta p_a$ yields

$$\pi_\gamma = \frac{k_{\theta\gamma}}{\frac{k_{\theta\theta}}{k_\theta} - \frac{p_{\theta\theta}}{p_\theta} + \frac{p_{a\theta}}{p_a}} - k_\gamma \tag{A9}$$

which verifies the claim. ■

Appendix 3

Claim 7 *The example defined by (11),(12) and (13) statisfies (10).*

Proof. In order to verify that condition (13) holds, we compute the relevant elasticities and substitute in (10). Hence, for all $f > 0$ we have:

$$\frac{\varepsilon^k_{\gamma\theta}}{\varepsilon^k_{\theta\theta} + \varepsilon^p_{a\theta} - \varepsilon^p_{\theta\theta}} = \frac{2(\gamma\theta + f)}{2\gamma\theta + f} > 1. \tag{A10}$$

■