



**HAL**  
open science

## Does Monitoring Decrease Work Effort?

David Dickinson, Marie Claire Villeval

► **To cite this version:**

David Dickinson, Marie Claire Villeval. Does Monitoring Decrease Work Effort?. *Games and Economic Behavior*, 2008, 63 (1), pp. 56-76. halshs-00276284

**HAL Id: halshs-00276284**

**<https://shs.hal.science/halshs-00276284>**

Submitted on 29 Apr 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Does Monitoring Decrease Work Effort?

## The Complementarity Between Agency and Crowding-Out Theories

David Dickinson <sup>a</sup>

Marie-Claire Villeval <sup>b</sup>

<sup>a</sup> Corresponding author

Appalachian State University, Dept. of Economics, Boone, NC 28608, U.S.A.  
phone 828-262-7652, fax 828-262-6105, Email: [dickinsond1@appstate.edu](mailto:dickinsond1@appstate.edu)

<sup>b</sup> CNRS, GATE, Ecully, F-69130; University of Lyon, Lyon, F-69007; University of Lyon 2, Lyon, F-69007; University of Lyon 1, Lyon, F-69003; ENS LSH, Lyon, F-69007; IZA, Bonn, D-53113. Email : [villeval@gate.cnrs.fr](mailto:villeval@gate.cnrs.fr)

**Acknowledgements:** The authors thank participants at the European Association of Labour Economists conference in Lisbon, and anonymous referees for extremely valuable comments and suggestions. We also thank Romain Zeiliger, who programmed the experimental environment. Funding from a grant from the ACI-INTERSHS program at the French ministry of research is gratefully acknowledged.

**Abstract**

Agency theory assumes that tighter monitoring by the principal should motivate agents to increase their effort, whereas the “crowding-out” literature suggests that the opposite may occur. These two assertions are not necessarily contradictory provided that the nature of the employment relationship is taken into account (Frey 1993). Results from controlled laboratory experiments show that many principals engage in costly monitoring, and most agents react to the disciplining effect of monitoring by increasing effort. However, we also find some evidence that effort is crowded out when monitoring is above a certain threshold. We identify that both interpersonal principal/agent links and concerns for the distribution of output payoff are important for the emergence of this crowding-out effect.

*JEL Code:* M5, J24, C92

*Keywords:* Principal-agent theory, Monitoring, Crowding-out, Motivation, Real effort experiment

## 1. INTRODUCTION

In an agency relationship, a principal may monitor agents' activity in order to increase effort. Assuming disutility of effort, a self-interested agent will work harder to reduce the probability of a sanction if caught shirking (Alchian and Demsetz, 1972; Calvo and Wellisz, 1978; Fama and Jensen, 1983; Laffont and Martimort, 2002; Prendergast, 1999). An alternative view from social psychology—'the crowding-out' hypothesis—considers that effort may be reduced by increased monitoring (see Frey, 1993).<sup>1</sup> Under this hypothesis, monitoring may reduce the intrinsic motivation of the task, which would reduce the agents' self-esteem or self-determination, and it could also be interpreted as an expression of distrust that violates a norm of reciprocity.<sup>2</sup> In other words, there are hidden costs of monitoring and sanctions that may dominate other incentives in certain cases. Frey (1993) argues that the more interpersonal is the principal/agent relationship, the more likely it is that more intense but still imperfect monitoring will backfire on the principal and actually reduce agent effort.

The aim of this paper is to use a controlled laboratory experiment to analyze the influence of the nature of the employment relationship on the relative importance of the disciplining and crowding-out effects of monitoring on agents' effort. To explore this principal-agent issue, we consider an employer/worker type relationship where the worker

---

<sup>1</sup> The crowding-out theory has been mostly developed by social psychologists in connection with the so-called cognitive evaluation theory (Deci, 1971; Deci et al., 1999). These analyses have notably emphasized the hidden cost of rewards (Lepper and Greene, 1978). The existence of a crowding-out effect of intrinsic motivation is, however, contested (Prendergast, 1999) or neglected by most economists. An early exception is Titmuss (1970) who argued that paying for blood donation would destroy the willingness to donate. See the related economic literature in Section 2.

<sup>2</sup> Corroborating this analysis, Bewley (1999) shows, from interviews with managers and labor leaders, that the risk for managers of using threats is a loss of worker initiative.

(i.e., the agent) engages in a real-effort task. Compared with much of the existing literature, this effort task is more likely to contain an intrinsic motivation component that could be affected by the experimental employer’s monitoring decision. An employer can monitor a worker by choosing the probability with which the worker’s output is audited. The worker then performs the task after being informed about the employer’s monitoring choice. Our main treatment variable, *Distant* or *Interpersonal*, is aimed at testing the Frey (1993) hypothesis. We implement the *Distant* relationship treatment by matching subjects as strangers in each round and preserving the anonymity of the pairs. We proxy a more *Interpersonal* relationship by utilizing a partner matching protocol (i.e., same subject pairing for ten rounds), *and* by removing subject anonymity. Indeed, in this treatment, we introduce subjects to one-another and allow each principal-agent pair to engage in five-minutes of face-to-face social interaction—discussion related to the experiment was prohibited.<sup>3</sup>

Our results indicate that the disciplining effect of monitoring predicted by agency theory dominates in both our distant and interpersonal relationship treatments. However, we also find evidence in the interpersonal treatment that increased monitoring crowds-out effort. Furthermore, because we include an additional treatment where the employer’s payoff is a continuously increasing direct function of worker output (*Variable*), as well as one where it is not (*Fixed*), we can explore the mechanism by which the crowding-out of

---

<sup>3</sup> While our *Interpersonal* treatment alters both anonymity and repetition with the same counterpart (relative to the *Distant* treatment), we feel that both are important components of ‘interpersonal’ relationships (i.e., it is harder to imagine an interpersonal one-shot relationship). An alternative experimental design might be considered if one is interested in identifying finer distinctions such as reputation versus social distance effects. We do, however, include gender variables in the econometric analysis of Section 5 to partly identify an effect that is due to lack of anonymity in the *Interpersonal* treatment, independent of the repeated interaction.

effort occurs. For example, if monitoring crowds-out effort by reducing intrinsic motivation, we should observe no difference in output between the two treatments, other things equal. The data indicate that crowding-out of effort occurs in the interpersonal employment relationship *only* when employer payoffs are a continuously increasing and direct function of worker output. This suggests that the mechanism by which crowding-out works is one of reciprocity, not intrinsic motivation reduction (even though we provide evidence of an intrinsic motivation for our experimental task).

Employers appear somewhat aware of the importance of reciprocity, but the reduction in monitoring intensity when their payoff depends primarily on worker output is not statistically significant. This is consistent with the hypothesis that principals, in general, behave more as predicted by agency theory than by the crowding-out theory. A failure to recognize that there may be implicit costs of monitoring may harm the principals' payoffs, however, as our results show that average payoffs are highest for those principals who monitor the least, and lowest for those who monitor the most. The results from our subjects provide evidence that crowding-out of effort should be a concern particularly when reciprocity can be inferred and relationships are more personal.<sup>4</sup>

This study contributes to the existing literature in three ways. First, the task we use in our experimental environment is more likely to have some intrinsic motivation compared to a simple allocation decision, which is important in the design if the reduction of intrinsic

---

<sup>4</sup> Houser, et al. (2007) note the importance of "intentions" on the effect of punishment threats. They compare a punishment threatened randomly by nature to an intentional punishment by another subject. We opt against a monitor-by-nature treatment in our research, but in some sense one could claim that a punish-by-nature treatment is a distant or abstract relationship. However, there may be an important, and uncontrolled, psychological distinction between a distant principal-agent and a *Nature*-agent relationship, which is really not a proper relationship. So, our design is more in line with the intent of Frey (1993).

motivation is a mechanism by which crowding-out occurs.<sup>5</sup> Secondly, we include *Interpersonal/Distant* treatments that alter the interpersonal nature of the employment relationship, which is hypothesized to be a key determinant of crowding-out according to Frey (1993). Finally, our *Fixed/Variable* treatment manipulation alters the link between worker output and employer payoff, which allows us to explore the mechanism by which crowding-out may occur in our environment—reciprocity or intrinsic motivation.

The remainder of this paper is organized as follows. Section 2 introduces the related literature. Sections 3 and 4 present the monitoring game and the experimental design and procedures. Section 5 analyzes the results and Section 6 concludes.

## 2. RELATED RESEARCH

Psychologists have provided compelling evidence that rewards can crowd-out intrinsic motivation (see the meta-analysis by Deci et al., 1999). Economists became recently interested in this issue (Drago, 1989; Frey and Jegen, 2001; Kreps, 1997; Benabou and Tirole, 2003) and empirical (mostly experimental) evidence has been collected on this ambiguous effect of rewards (Bohnet et al., 2000; Fehr and Gächter, 2002; Frey and Oberholzer-Gee, 1997; Gneezy and Rustichini, 2000a).<sup>6</sup>

---

<sup>5</sup> Though our effort task is less comparable to one that would occur in a natural work setting, it adds to the literature of effort experiments that do not entirely induce the effort environment. Such environments are more externally valid, though potentially less internally controlled from an experimental standpoint.

<sup>6</sup> Kreps (1997) states that if employees develop intrinsic motivation in reaction to fuzzy extrinsic incentives, the introduction of explicit incentives may diminish intrinsic motivation. Benabou and Tirole (2002) try to reconcile psychological and economic approaches showing that rewards may be strong or negative reinforcers depending notably on the ability of the agent, and the asymmetry of information regarding the agent's talent and the nature of the task. Gneezy and Rustichini (2000a) show that the subjects working for free reached a much higher outcome than those who were paid a small amount of money. Fehr and Gächter (2002) state that

When studying explicit *punishment* incentives, experimental research also indicates a potentially detrimental effect of sanctions (e.g., Gneezy and Rustichini, 2000b; Lin and Yang, 2006; Fehr and Schmidt, 2000; Fehr and Rockenbach, 2003; Fehr and List, 2004). The negative effect of sanctions is due to a breach of the social norm of trust or to the reduction of the psychological cost arising from the violation of a norm. For their part, Houser et al. (2007) show that threats of sanctions *increase* return amounts in an investment game—they conclude this even when a sanctioning threat is intentional and the exchange environment allows for reciprocity. But this is conditioned by the credibility of threats. With a different game, Gneezy (2003) shows that the effect of sanctions and rewards on performance may be either positive or negative, depending on the size of the incentives.

However, evidence of effort crowding-out due to uncertain monitoring, which is a probabilistic sanction, remains scarce and the lack of naturally-occurring data usually leads to indirect tests of competing hypotheses. Some supporting empirical evidence for the crowding-out hypothesis of monitoring is found in a survey of managers, where Barkema (1995) and Frey (1993) document a negative effect of monitoring on hours worked when the principal is a CEO but a positive effect when the manager is supervised impersonally by a parent company. A field experiment on call centers shows that employees do not respond to the exogenous manipulation of monitoring rates when they think the employer is being fair (Nagin et al., 2002).

---

incentive contracts are less efficient than contracts without any incentives because of the crowding-out effect. Nevertheless, principals use incentive contracts because they can obtain a higher share of the surplus.



Nevertheless, field and survey data do not control for many potentially confounding factors. A laboratory study can help mitigate such concerns by collecting controlled data on both monitoring and performance. A proper experimental design must control variables such as the fear of dismissal, the search for co-workers' esteem, or the disapproval of the principal, all of which can generate effects on effort and performance that might be mis-attributed to intrinsic motivation in field data (Kreps, 1997). The limited number of controlled laboratory studies in this area find some evidence to support a negative effect of monitoring on performance (e.g., Schulze and Frank, 2003; Guerra, 2002). The restriction of autonomy represents a different type of control of workers' activity. Falk and Kosfeld (2006) allow principals to restrict the agents' choice set and find that, when 'controlled' in this way, agents are less generous towards the principal as a response to the signal of distrust created by the restriction of possible actions. None of these studies, however, control for the nature of the relationship between principals and agents.

### **3. THE MONITORING GAME**

The game involves two players, a principal and an agent. Output,  $y$ , is produced by the agent and depends on agent effort and task difficulty. In the *Variable* treatment, the principal's payoff directly and continuously increases in the agent's output. Alternatively, in the *Fixed* treatment, the principal's payoff is only affected discontinuously by agent's effort through a partial recovery of the shirking penalty. Thus, the direct and continuous

link between agent effort and principal payoff is severed in the *Fixed* treatment, and the principal's dominant compensation is a fixed payoff.<sup>7</sup>

**The Variable Treatment.** In the first stage of the game, the principal offers a contract  $(W, w, \underline{w}, \hat{y}, \underline{y}, m)$  to the agent, consisting of three wage levels,  $W$ ,  $w$  and  $\underline{w}$ , corresponding to various levels of outputs, a desired level of output  $\hat{y}$ , a minimum output requirement  $\underline{y}$ , and a probability of audit  $m$  (i.e., the monitoring intensity). For simplicity, we fix  $\underline{y} = 40$  and  $\hat{y} = 75$  throughout. Payment to the agent is  $W = 100$  if his output is not audited or audited and  $y \geq \hat{y}$ . If audited and  $\underline{y} \leq y \leq \hat{y}$  then  $w = 60$  is paid, whereas if  $y \leq \underline{y}$ , then the agent receives the minimum wage  $\underline{w} = 20$ . The monitoring intensity,  $m \in [0, .1, \dots, 1.0]$ , represents the probability that the agent's output will be audited, and it is the principal's only decision variable. Monitoring is costly for the principal beyond some threshold  $\underline{m}$ , with  $\underline{m} = .2$ . The intuition is that below a certain level, monitoring is not costly because simple observation may be sometimes sufficient to detect shirking. If shirking at some level has been verified, the principal punishes him by paying a lower wage,  $w$  or  $\underline{w}$ . The differences between  $W$  and  $w$  and between  $W$  and  $\underline{w}$  can be considered as fines, and the size of the fine depends on the extent of shirking.

The principal chooses monitoring intensity according to an induced cost function:

---

<sup>7</sup> As can be seen in equations (2) and (4) below, there is a partial 're-capture' of the shirking penalty by the principal. The payoff link between agent effort and principal effort is, however, discontinuous (and also dominated by the direct agent output payoff link in the *Variable* treatment). We show in an appendix (available on request) that our parameterization is such that the theoretical predictions for rational and selfish principals and agents are distinct. Thus, we have not inadvertently created an indirect payoff effect in (2) and (4) that alters the key theoretical prediction that monitoring and effort are both higher in the *Variable* treatment

$$c(m) = \frac{(1.5m - \underline{m})^2}{b} \quad \text{for } m \geq .3 \quad (1)$$

with  $\underline{m} = .2$ , the threshold for free monitoring, and  $b = .02$ . After rounding numbers, monitoring costs are displayed in Table 1. The (expected) payoff function of the principal in the *Variable* treatment is given by:

$$\pi_{PV} = \begin{cases} vy_i - W - c(m) & \text{if } y_i \geq \hat{y} \\ vy_i - W + m(W - w) - c(m) & \text{if } \underline{y} \leq y_i \leq \hat{y} \\ vy_i - W + m(W - \underline{w}) - c(m) & \text{if } y_i \leq \underline{y} \end{cases} \quad (2)$$

where  $y_i$  refers to the output of the  $i^{\text{th}}$  agent and with  $v = 2.5$  (each unit of output of the agent increases the principal's payoff by 2.5). The second and third equations in (2) specify the re-capture of the shirking penalty.

In the second stage, the agent is informed of the contract  $(W, w, \underline{w}, \hat{y}, \underline{y}, m)$ .<sup>8</sup> Knowing his payoff function, the payoff function of the principal, the outcome requirement, and the monitoring probability, the agent then chooses effort. The agent's task, which consists of progressing along a curve where output is measured by the height reached on this curve, is described in more detail in the next section. In short, progression on the curve is made by means of regular-steps that are free and by means of rapid-steps that are costly. Performing a higher output usually entails a marginally increasing monetary cost that is linked to speed of progression chosen by the agent. In addition,  $c(0)=0$  and  $c(\underline{y})=0$ . Indeed, curves have been designed such that the minimum required output can always be performed without any monetary cost, for all possible draws from the

---

<sup>8</sup> There is no stage in the game where the agent accepts or rejects the contracts. The experiment is setup so that the participation and incentive compatible constraints are always met from the beginning. A subject who would like to express negative reciprocity must do it through his effort decision.

distribution. In contrast, it is common information that on average, it costs 20 points to achieve the required output  $\hat{y}$ .  $W \geq c(y^*)$  ensures that the participation constraint holds.

The  $i^{\text{th}}$  agent's (expected) payoff function is given by:

$$\pi_i = \begin{cases} W - c(y_i) & \text{if } y_i \geq \hat{y} \\ W - c(y_i) - m(W - w) & \text{if } \underline{y} \leq y_i \leq \hat{y} \\ W - c(y_i) - m(W - \underline{w}) & \text{if } y_i < \underline{y} \end{cases} \quad (3)$$

In stage three, once the agent has performed the task, the monitoring probability is applied, the audit occurs or not depending on the draw, and the payoffs are displayed.

**The Fixed Treatment.** This treatment is similar to the *Variable* treatment, except that the principal's payoff is fixed and no longer increasing in the agent's output. We now have:

$$\pi_{PV} = \begin{cases} f - W - c(m) & \text{if } y_i \geq \hat{y} \\ f - W + m(W - w) - c(m) & \text{if } \underline{y} \leq y_i \leq \hat{y} \\ f - W + m(W - \underline{w}) - c(m) & \text{if } y_i \leq \underline{y} \end{cases} \quad (4)$$

where  $f = 180$  is the flat fee given to the principal. As noted before, the reduced wage in the event of low audited output accrues as a benefit to the principal in (4). Because this recapture effect is identical across treatments, the key distinction in (2) and (4) is the direct connection between agent output and principal payoff.

The parameterization we use is such that the selfish predictions of this model are higher monitoring and output in the *Variable* treatment ( $m_v^* = .5$ ,  $y_v^* = 75$ ), compared to the *Fixed* treatment ( $m_F^* = .3$ ,  $y_F^* = 40$ ).<sup>9</sup> However, these predictions do not account for behavioral motivations resulting from *implicit* incentives. It may be tempting for a trustful

---

<sup>9</sup> These theoretical predictions are available in an additional Appendix, on request.

principal to choose a positive but minimal degree of monitoring since this costs nothing (monitoring is free below  $m=.3$ ). Therefore, a principal who chooses  $m_y = 0$  signals an even greater degree of trust. In these circumstances, the agent might be willing to reciprocate by increasing his effort and thus increasing the principal payoff in the *Variable* treatment. Similarly, behaviorally we hypothesize that reciprocity is decreasing in  $m$ . Indeed, the higher the monitoring intensity, the lower the trust of the principal in her agent. In the *Fixed* treatment, an agent who produces output levels other than  $y=0, 40, \text{ or } 75$  might be expressing a desire to simply perform the task (i.e., intrinsic motivation). The overall pattern of behavior across all treatments does not support the conclusion that agents are confused.

#### 4. EXPERIMENTAL DESIGN AND PROCEDURES

**Design of the task.** The design of the agent's task is driven by our desire to balance the need for experimental control with the need for a somewhat intrinsically motivating task. The task, first used in Montmarquette et al., (2005), consists of the search of the highest value of a function in a two-dimension space defined vertically by height and horizontally by distance from the origin. The curve corresponding to this function is increasing, with a maximum of three plateaus. During a one-minute 'work' period, the agent-subject clicks a button to progressively uncover the curve. The subject moves by discrete horizontal steps, which may increase the vertical height of the curve (unless on a plateau). The curve and its surface only become visible as the subject progresses. The performance, i.e. the output achieved by the subject, in a period is given by the height reached on the curve, between 0

and 100, which depends notably on the number of moves. The subject screen indicates the remaining time in the round, the cumulative costs, and the height reached (see Figure 1).

The cost of effort is represented through the choice of the progression speed. It is impossible to reach the required height,  $\hat{y}=75$ , during the one-minute period by using the regular speed only. Two buttons are available: regular (“1-steps”) or fast (“2-steps”), the second option enables one to move twice as rapidly. The agent can switch speeds whenever he wants. The regular speed is costless, whereas each 2-step is costly. To ensure cost convexity, each of the first ten 2-steps costs .4 point, each of the ten next 2-steps costs .6, and each 2-step after that costs 1 point. On average, reaching  $\hat{y}=75$  costs 20 points, which is common knowledge, with a minimum cost of 0 and a maximum of 70.

Each new period is associated with a new randomly chosen curve to maintain a cognitive interest in the task and to limit the possibility that subjects can memorize the particular profile of a curve and exploit that in making his effort choice.<sup>10</sup> The cognitive dimension of this task relates to the uncertainty about the shape of the curve, to time pressure and to the decision to use the fast speed. Without any *ex ante* information about the shape of the curve, paying each additional costly step requires a continuous trade-off between its marginal cost and its expected marginal benefit considering the time left.

**Design of the employment relationship.** All the subjects were administered both the *Distant* and *Interpersonal* treatments to enable within-subject comparisons. Specifically, a stranger matching protocol is used to proxy a distant employment relationship with no

---

<sup>10</sup>The set of curves used in this experiment can be found at the following address: [http://koala.gate.cnrs.fr/groupware/g\\_gate/Courbes-MCV-DI.doc](http://koala.gate.cnrs.fr/groupware/g_gate/Courbes-MCV-DI.doc)

repeated interactions. Principals and agents are randomly re-matched after each period and this is common knowledge.

To proxy an interpersonal relationship, the *Interpersonal* treatment allows for social exchange by using a partner matching protocol and by lifting anonymity within pairs. Specifically, a principal-agent pair is randomly matched and remains paired for ten periods. In addition, the subjects of each pair are introduced to each other. They are seated face-to-face and they are given five minutes to talk together for introducing themselves to each other. They are not allowed to comment on the experiment or their decisions, or to pass any side-payment agreement. To facilitate this introduction and to focus the discussions on personal issues, we distributed a questionnaire sheet to each pair.<sup>11</sup> At the end of this introduction process, the members of a pair remain seated side by side but no longer face-to-face and they are not allowed to communicate any more. This proximity is such that subjects can anticipate that at the end of the session the partners will be able to criticize their respective decisions. This design may have behavioral consequences by narrowing the social distance and by favoring empathy between the principal and the agent.<sup>12</sup>

**Experimental procedures.** Ten 20-period experimental sessions were conducted at the experimental laboratory of GATE (Groupe d'Analyse et de Théorie Economique) in Lyon,

---

<sup>11</sup> This questionnaire consisted of two parts, one about each member. Each principal was requested to ask the questions to her agent and to write her answers, and vice-versa for the agent. It included questions about their given name, their number of siblings, their favorite music group, their hobby and the location of their past summer vacations.

<sup>12</sup> The nature of our *Interpersonal Relationship* treatment may be criticized as not seriously creating anything interpersonal. While we do not contend that our five-minutes of face-to-face social interaction created any life-long relationships, it was an extra step towards removing the social distance that has been found significant in other experimental environments (on the impact of personal identification, see Kachelmeyer and Shebata, 1997; Sally, 1995; Charness et al., 2007; Falk et al, 1999). This time of social interaction was quite popular among the subjects, and so we feel confident that we created a meaningful distinction in interactions between the anonymous *Distant* relationship treatment and the *Interpersonal* treatment.

France, using the Regate software (Zeiliger, 2000). Five sessions implemented the *Variable* Treatment and five other sessions the *Fixed* Treatment. For both treatments, three out of five sessions implemented first the stranger matching protocol and then the partner protocol, the two other sessions implemented the reverse order.<sup>13</sup> The participants were 182 undergraduate students mainly from engineering and business schools.

On average, a session lasted 75 minutes. Transactions were conducted in Experimental Currency Units, with the ECU-Euro conversion rate set at  $\text{ECU } 150 = \text{€ } 1$ . The final payoff was given by the sum of the earnings in each period. In addition, a show-up fee consisted of two elements. At the beginning of the session and before the presentation of the instructions, each subject had to choose between either participating in a lottery with an expected payoff of € 2.5 or taking a certain gain of € 2, at the end of the session. This gives an (imperfect) indication of the subjects' risk aversion. In addition, because losses are possible, each subject received an initial endowment of 150 ECUs. The average earnings were € 13.42 (S.D.=2.65). Subjects were immediately paid in cash in a separate room.

Participants were randomly assigned to a computer terminal. Instructions (phrased in neutral terms) were distributed to participants and read aloud by the experimenter (see Appendix). Questions were answered privately. Two practice periods were run by all the subjects<sup>14</sup> and a questionnaire was passed to check the understanding of the instructions.

---

<sup>13</sup> Eight sessions were initially realized in which the order of the curves was the same in both parts of the experiment. Two additional control sessions (one for each treatment) were conducted in which the order of the curves was changed in the second part of a session to even further control for possible memory effects. A means test shows that average effort is not significantly affected by the order of the curves, at the 5% level.

<sup>14</sup> At this stage the subjects are not aware yet of the role they will be assigned, and it makes the future principals aware of the possible difficulty of the task.



Each subject was then randomly assigned his role by the computer, which remained constant throughout the session. At the end of each period, both the principal and the agent receive feedback on the output realized, whether audit occurred, and their actual payoffs. The instructions for the second part were distributed at the end of the first part.

## 5. EXPERIMENTAL RESULTS

Before analyzing the determinants of principals' and agents' behavior, we examine descriptive statistics regarding the average monitoring intensity and output. These descriptive statistics for all treatments are shown in Table 2.

We find that, on average, mean monitoring levels are lower in the *Fixed* treatments than in the *Variable* ones, as predicted by the theory based on selfish motivations, but mean agent output levels are *not* lower in the *Fixed* treatments, contrary to the predictions. The data can be organized into 5 perfectly independent observations by averaging outcomes for each session. Using these units of observation, the signed-rank test indicates that monitoring levels are higher than predicted in the *Fixed* treatments (i.e., reject  $H_0: m=m^*=.3$  in favor of one-sided alternative), but lower than predicted in the *Variable* treatments (rejecting  $H_0: m=m^*=.5$ ) ( $p=.05$  for all results). However, a comparison of monitoring levels across treatments using the Mann-Whitney test indicates that they are significantly lower ( $p=.05$ ) in the *Fixed-Interpersonal* treatment than in both the *Variable-Distant* and *Variable-Interpersonal* treatments. This is visible in Figure 2 that displays the cumulative density function of monitoring decisions.

For the agents, mean output levels are significantly lower than predicted in the *Variable* treatments (rejecting  $H_0: y=y^*=75$  in favor of the one-sided alternative) and significantly higher than predicted in the *Fixed* treatments (rejecting  $H_0: y=y^*=40$ ) (sign-rank tests,  $p=.05$  in all cases). Agent output shows some significant differences in comparisons across treatments. In both the *Variable-Distant* and *Variable-Interpersonal* treatments, the mean output is significantly lower than mean output in the *Fixed-Interpersonal* treatment (Mann-Whitney,  $p=.05$  in both cases). These results give an initial indication that the higher monitoring by the principals in the *Variable* treatments does not necessarily lead to higher agent output, which is consistent with some level of effort crowding-out.<sup>15</sup> Before examining the link by which crowding-out may occur, it is important to highlight the evidence for intrinsic motivation in the effort task.

**Intrinsic motivation.** The data provide evidence that intrinsic motivation exists for this task. First, we show that a significant proportion of subjects outperform the standard. Specifically, across the treatments (*VDR=Variable-Distant Relationship*, *VIR=Variable-Interpersonal Relationship*, *FDR=Fixed-Distant Relationship*, and *FIR=Fixed-Interpersonal Relationship*), the number of subjects who achieve output greater than the required level of 75 is 32%, 29%, 30%, 36%, respectively. In the *Variable* treatments, this could result from a desire to reciprocate or reward the principal. However, in the *Fixed* treatments, this can only be explained by an intrinsic desire to complete the costly task. An

---

<sup>15</sup> Similar tests can be performed using the subject's average outcome in a treatment as the unit of observation. Though these are not clearly independent observations as with the session-level data, the results of such individual-level tests of principal monitoring choices concur with the session-level tests. Monitoring levels in either of the *Fixed* treatments are lower than monitoring levels in either of the *Variable* treatments ( $p<.05$  for all pairwise tests). For the agents, output levels are not significantly different ( $p>.05$ ) when using this alternative unit of observation.

output  $y > 75$  is also beyond that which would result from a desire to respect the norm of achieving the benchmark output level  $y = 75$ .<sup>16</sup> Secondly, we observe that people exert an effort even when they know they cannot be monitored. Though a monitoring choice of  $m = 0$  was relatively rare, we observe that 93% (N=14), 100% (N=13), 100% (N=7), and 95% (N=19) of the subject rounds in *VDR*, *VIR*, *FDR*, and *FIR*, respectively, produced  $y > 0$  when  $m = 0$ —a full 83% (44 of 53 across all treatments) produced  $y > 40$  when  $m = 0$ . Moreover, of those who produced  $y > 40$  when  $m = 0$ , 59% (26/44) sacrificed money to do so, including 48% (10/21) in the *Fixed* treatments where such choices had no effect on principal payoffs. In other words, there seems to be enough evidence of intrinsic task motivation to explore its role in effort crowding-out.

Summary statistics do not control for time, task difficulty, or individual-specific effects. For these reasons, we also conduct a more formal econometric analysis of the monitoring and effort choices of principals and agents, respectively.

**Principals' monitoring decisions.** To explain the monitoring intensity chosen by the principals, we estimate separate random-effects GLS models of the *Variable* versus *Fixed* treatments. We model the monitoring choice as a function of the agents' past behavior (captured by lagged output). We control for the nature of the employment relationship by including the *Interpersonal* treatment variable and an interacted variable (*Interpersonal*\*lagged output). We also include a time trend and dummy variables for

---

<sup>16</sup> From an extensive U.S. survey about worker motivations, Minkler (2004) shows that the rank-order of motivations is moral duty, intrinsic motivation, peer pressure and incentives. To a question about their motivation if it is impossible for their employer to check up on them, 83% of the respondents answered that they were "very likely" to work hard, and 12% that they were "somewhat likely" to do so. Our data reveal a smaller proportion but we directly measure the work output and not only declarations of intention.

periods 1 and 11 to take into account possible learning and restart effects. We include a dummy variable for the order of the *Interpersonal* versus *Distant Relationship* treatments. We add gender-based variables to control for the potential effects of the principal's gender, as well as an interaction variable for a same-gender pair in the *Interpersonal* treatment. The random effects error-term structure accounts for the potential interdependence of the error term for a particular principal across her twenty rounds of observations.<sup>17</sup> Results are reported in Table 3.

Crowding-out is principally a hypothesis relating to agents' behavior, but we predict that principals' behavior should be affected if they anticipate a crowding-out effect of monitoring. In the *Variable* treatment, the pure effect on principals of an interpersonal relationship is to reduce monitoring. Because we do not estimate a similar tendency in the *Fixed* treatment, these results might suggest that principals in interpersonal relationships monitor less heavily when reciprocity is a concern. This is consistent with the hypothesis that principals view crowding-out as a phenomenon related to reciprocity, rather than intrinsic motivation of the task. At the same time, interestingly, we also find that there is a tendency to monitor more strictly over time in the *Variable* treatment, whereas the tendency is the opposite for the *Fixed* treatment where agent output does not affect the employer's payoff. This may give an indication that the principal's decisions are, over time, trending towards equilibrium, *ceteris paribus* (i.e., they may be learning to play the theoretical prediction). We should therefore expect more divergence in the monitoring intensity

---

<sup>17</sup> The results are similar in terms of magnitude and statistical significance with a fixed-effects modeling. Random effects results are more appropriate if one believes that the cross-sectional units (i.e., the subjects) were randomly drawn from a population of decision-makers. A tobit model was also explored, but there was difficulty in attaining proper model convergence. However, the limit monitoring probabilities were only chosen in 3-4% of the observations, and so the model we report is reasonable.

between the two treatments over time. Another result is that the monitoring intensity is decreasing in lagged agent output in all treatments, even when pairs are rematched after each period.

Interestingly, this regression shows that in the *Variable* treatment, where reciprocity may be expected to influence the principals' payoffs, females monitor significantly less heavily than males in general, and they monitor even less when they know they are faced with a male agent. In contrast, in the *Fixed* treatment, there is no gender difference in the principals' behavior and the principals of the same-gender pairs tend to monitor less compared to mixed-gender subject pairs. These gender results are important, because they are an indication of how lack of anonymity, as well as repeated play with the same counterpart, are both important outcome determinants in the *Interpersonal* treatment.

**Agent response to monitoring.** To test the crowding out hypothesis, we again use a random-effects GLS specification to model agent output as a function of the monitoring intensity, which the agent knows prior to making his output decision. We add explanatory variables to control for the difficulty of the task<sup>18</sup> and individual characteristics. Table 4 displays the results of these estimations.

In evaluating the crowding-out hypothesis, we first note that agent output is significantly higher in an interpersonal relationship, but only in the *Variable* treatment. This is consistent with our results from Table 3. Controlling for the task difficulty and

---

<sup>18</sup> The degree of difficulty depends on the horizontal location of the 75-height and the maximum height of 100. An index of difficulty is calculated as  $10(D_{75} + \sqrt{(D_{100} - D_{75})})$ , where  $D_{75}$  is the abscissa at the origin of the 75-height and  $D_{100}$  is the abscissa of the maximum height. It varies in between 1549 and 1868).

individual characteristics,<sup>19</sup> we obtain a significant and independent agency theory effect of monitoring on agent output—coefficients on *Monitoring Intensity* are significant and positive in both the *Variable* and the *Fixed* treatments: *ceteris paribus*, the higher the monitoring intensity, the higher the agent's outcome. Thus, in Table 4 we have clear evidence that extrinsic incentives matter.

However, in the *Variable* treatment, we observe both a significant downward output trend over time and a significant negative interaction effect between the *Monitoring Intensity* and *Interpersonal Relationship*. Both elements contribute to identifying a crowding-out effect of effort. The first element indicates that the impact of extrinsic incentives is decreasing over time in this environment since the downward output trend goes along with an upward trend of monitoring intensity across rounds. The second element indicates that there is a negative impact of monitoring on performance when subjects are engaged in an interpersonal relationship with no anonymity and repeated interactions. In contrast, as behaviorally predicted, there is no evidence of crowding-out when the principal-agent relationship is distant.

In addition, we find that the interaction between the monitoring intensity and the interpersonal relationship is not significant in the *Fixed* treatment. Because the crowding out effect of monitoring only occurs in the *Variable* treatment and not in the *Fixed* treatment, this is consistent with our hypothesis that crowding out functions through reciprocity. Recall that an agent can only truly reciprocate intense monitoring with low

---

<sup>19</sup> The coefficient estimates indicate that a more difficult task lowers agent output nonlinearly. We also observe a significant start or 're-start' effect that increases output in rounds 1 and 11. In contrast, neither risk attitude nor gender variables are significant predictors. Together, these results support the hypothesis that more speculation occurs with principals, who face the more uncertain choice task.

effort/output in the *Variable* treatment where agent output affects principal payoffs in a continuous and direct fashion. If crowding-out were to function through a decrease of intrinsic motivation, we should observe the same significant negative effect of monitoring on performance in an interpersonal relationship in the *Fixed* treatment, which we do not.

As a consequence, there are hidden costs of monitoring that can have real payoff implications. Most of our principals chose average monitoring levels between  $m=.3$  and  $m=.5$  ( $N=154$  observations). These principals averaged payoffs of 76.7 ECUs in each round. Principals that averaged low monitoring levels — $m<.2$ — across the 10 rounds of a treatment ( $N=11$ ) had significantly higher average payoffs of 81 ECUs per round ( $p=.02$  for the signed-rank test, favoring the alternative hypothesis that these average payoffs were greater than 76.7 ECUs). Principals that averaged high monitoring levels — $m>.5$ — for the treatment ( $N=17$ ) had significantly lower earnings of 54.5 ECUs per round ( $p=.01$  for the signed-rank test).

In sum, the key results from our data show evidence that monitoring exerts a disciplining effect but also that more intense monitoring crowds out effort, and this occurs when the principal-agent relationship is more personal, which is consistent with Frey (1993). Monitoring may thus generate hidden costs to the principals. Furthermore, our *Variable/Fixed* payment treatments help identify that the mechanism by which crowding out occurs in our experiment is reciprocity, as opposed to reduction of the task's intrinsic motivation.

## 6. CONCLUSION

Agency theory predicts that positive (rewards) and negative (monitoring and sanctions) incentives support the agent's level of effort in equilibrium. However, hidden costs of rewards have been well-documented in the literature, especially in social psychology and, more recently, in behavioral economics. Empirical evidence for the crowding-out effect of monitoring has been scarce. While a recent and small literature has examined the potential negative effects of certain control devices, this paper is the first to explore whether endogenous monitoring probability—an uncertain punishment device—will crowd-out effort under conditions related to the nature of the principal-agent relationship. Thus, we attempt to fill an important gap in the literature on this topic.

The model provided by Frey (1993) states that monitoring by a principal primarily exerts a disciplining effect on agent's effort in distant employment relationships, whereas it has a crowding-out effect on effort when the principal and the agent are engaged in an interpersonal relationship. We provide a test of this hypothesis. We also test whether crowding-out of effort results from a reduced intrinsic motivation for performing the task or if it is due to an agent's desire to reciprocate a perceived mistrust from the principal. We use laboratory methods to generate more controlled data on principal-agent relationships than is available in the field, and we employ a task that is likely to contain intrinsic motivation for the subjects. We vary both the nature of the employment relationship and the way agent output influences principal payoffs. This latter treatment allows us to distinguish the mechanism by which crowding-out may function, which is another unique design feature of our experiments.

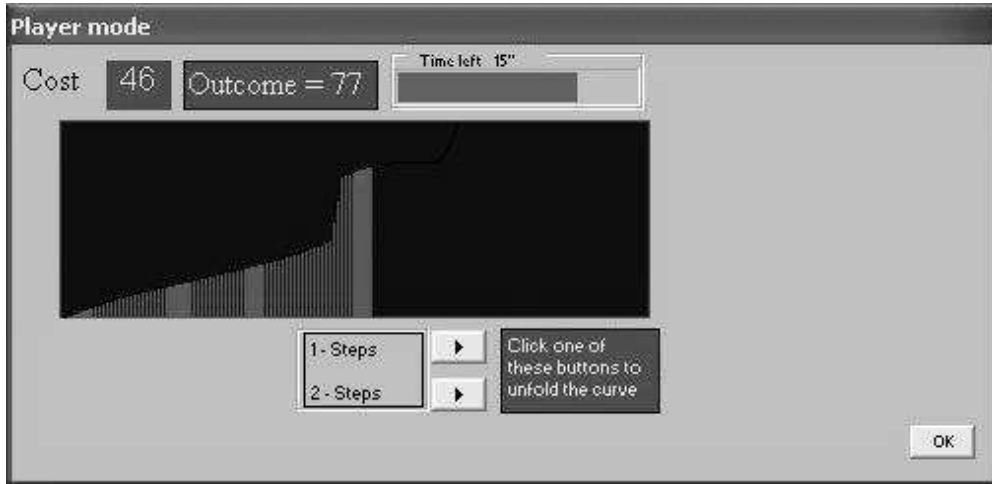


Our data provide ample evidence of the disciplining effects of monitoring, which is supportive of agency theory. However, we also find evidence that the output of our agents is crowded-out when an employer monitors more heavily in the interpersonal relationship treatment. Accordingly, monitoring may have hidden costs. Consistent with Frey's (1993) hypothesis, monitoring contains both positive and negative incentives from the principal's perspective. A key question is which effect—the disciplining or crowding out effect—dominates under which circumstances, and our evidence indicates that a reduced social distance between the principal and agent is a major condition for the crowding-out effect to emerge. But this is not all that matters. Our second treatment manipulation indicates that this crowding out effect occurs in more interpersonal relationships only when the principal's payoff directly and continuously increases in the agent's output. In other words, the mechanism by which crowding-out works, at least in our experimental environment, is through reciprocity. The mistrust the agent may feel when heavily monitored is reciprocated through lower output in the *Variable* treatment, while intense monitoring does not have this effect in the *Fixed* payoff treatments.

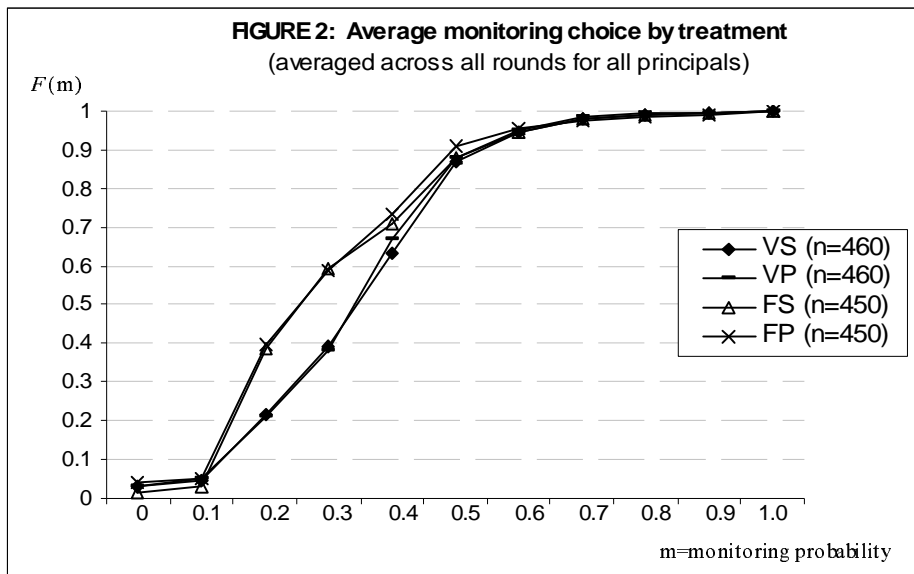
Our results must still be taken with caution. For example, our task is still quite abstract relative to a real-world labor task. We nonetheless provide evidence that it generated some intrinsic motivation. Other tasks, no doubt, will possess more intrinsic motivation than ours. In such cases, we would predict that crowding-out may be even more severe, because the avenue of intrinsic motivation is more apt to be disrupted. The fact that we do find significant behavioral differences in our contrived treatments strengthens our belief in the potentially overlooked consequences of monitoring in a significant number of

environments with limited social distance. From a theoretical point of view, our results also underline the importance of enriching agency theory by modeling incentives that preserve the complementarity of extrinsic and intrinsic motivation.

**FIGURE 1**  
Effort task example



**FIGURE 2**  
Distribution of the average monitoring levels by treatment



**Table 1.** Monitoring costs by intensity of monitoring

$m$	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
$c(m)$	0	0	0	3	8	15	24	36	50	66	84

**Table 2.** Mean monitoring and output levels by treatment

Treatment	Variable-Distant	Variable-Interpersonal	Fixed-Distant	Fixed-Interpersonal
Monitoring level	$\bar{m} = 3.90$	$\bar{m} = 3.87$	$\bar{m} = 3.48$	$\bar{m} = 3.37$
Agent output	$\bar{y} = 68.62$	$\bar{y} = 68.52$	$\bar{y} = 70.46$	$\bar{y} = 71.66$

Note: these values are averaged across rounds for each subject, then across all subjects.

**Table 3.** The determinants of monitoring intensity

Dependent variable: Monitoring intensity	Random effects GLS model	
	Treatment	Fixed
<i>Interpersonal</i> treatment	-.312* (.063)	.086 (.669)
Lagged output	-.018*** (.000)	-.014*** (.003)
<i>Interpersonal</i> * lagged output	.001 (.507)	.002 (.437)
Round	.034*** (.000)	-.046*** (.000)
First round of treatment (1 or 11)	-.098 (.513)	-.002 (.992)
<i>Interpersonal</i> before <i>Distant</i> treatment	-.012 (.965)	-.336 (.178)
Principal's risk aversion	.127 (.672)	.417* (.100)
<i>Male</i> principal	.570** (.043)	.241 (.325)
<i>Same Sex</i> * <i>Interpersonal</i> treatment	.294* (.079)	-.400** (.041)
Constant	4.393*** (.000)	4.789*** (.000)
N = number of observations	920	900
LM test for one-factor model	834.05***	286.73***
Model test $\chi^2_9$	66.61***	61.37***

Note: The regressions were performed using Limdep 8.0. P-values are in parentheses. \*, \*\*, \*\*\* indicate statistical significance at the .10, .05, and .01 level, respectively, for the two-tailed test.

**Table 4.** The determinants of agent output

Dependent variable: Level of output	Random-effects GLS model	
	Treatments	Fixed
<i>Interpersonal Relationship</i> treatment	3.89** (.045)	.847 (.614)
Monitoring intensity	1.117*** (.000)	1.192*** (.000)
<i>Interpersonal Relationship</i> * Monitoring intensity	-.895** (.041)	.378 (.340)
<i>Round</i>	-.130** (.036)	.090 (.155)
First round of treatment (1 or 11)	4.802*** (.000)	4.688*** (.000)
<i>Interpersonal Relationship</i> treatment before <i>Distant Relationship</i> treatment	.052 (.988)	1.883 (.231)
Agent's risk aversion	3.123 (.425)	1.358 (.435)
<i>Male</i> agent	1.615 (.601)	.566 (.713)
<i>Same Sex</i> * <i>Interpersonal Relationship</i>	-.630 (.630)	-1.754 (.162)
Task difficulty index	-.687*** (.000)	-.384*** (.005)
Task difficulty index squared	.0002*** (.000)	.0001*** (.006)
Constant	666.690*** (.000)	398.775*** (.001)
N = number of observations	920	900
LM test for one-factor model	2036.22***	271.93***
Model test $\chi^2_{11}$	65.40***	68.73***

Note: The regressions were performed using Limdep 8.0. P-values are in parentheses. \*, \*\*, \*\*\* indicate statistical significance at the .10, .05, and .01 level, respectively, for the two-tailed test.

**REFERENCES**

Alchian, A., Demsetz, H. 1972. Production, information costs, and economic organization. *Amer. Econ. Rev.* 62, 777-95.

Barkema, H., 1995. Do top managers work harder when they are monitored? *Kyklos*. 48, 19-42.

Benabou, R., Tirole, J. 2003. Intrinsic and extrinsic motivation. *Rev. Econ. Stud.* 70, 489-520.

Bewley, T. 1999. *Why Wages Don't Fall During a Recession?* Harvard University Press, Harvard.

Bohnet, I., Frey, B., Huck, S. 2000. More order with less law: On contract enforcement, trust and crowding. *Amer. Polit. Sci. Rev.* 95, 131-44.

Calvo, G., Wellisz, S. 1978. Supervision, loss of control, and the optimum size of the firm. *J. Polit. Economy*. 86, 943-52.

Charness, G., Rigotti, L., Rustichini, A. 2007. Individual behavior and group membership. *Amer. Econ. Rev.*, Forthcoming.

Deci, E. 1971. Effects of externally mediated rewards on intrinsic motivation. *J. Pers. and Social Psych.* 18, 105-15.

Deci, E., Koestner, R., Ryan, R. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psych. Bull.* 125, 627-68.

Drago, R. 1989. Supervision and high wages as competing incentives: A basis for labour segmentation theory. in R. Drago and R. Perlman, *Microeconomic Issues in Labour Economics: New Approaches*. Harvester Wheatsheaf, New York, 41-61.

Falk, A., Gächter, S., Kovacs, J. 1999. Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. *J. Econ. Psych.* 20, 251-84.

Falk, A., Kosfeld, M. 2006. The hidden cost of control. *Amer. Econ. Rev.* 96, 1611-30.

Fama, E., Jensen, M. 1983. Separation of ownership and control. *J. Law Econ.* 26, 301-51.

Fehr, E., Gächter, S. 2002. Do incentive contracts crowd-out voluntary cooperation? University of Zurich Working Paper, 34.

Fehr, E., List, J. 2004. The hidden costs and returns of incentives – trust and trustworthiness among CEOs. *J. Europ. Econ. Assoc.* 2, 743-71.

Fehr, E., Rockenbach, B. 2003. Detrimental effects of sanctions on human altruism. *Nature*. 422, 137-40.

Fehr, E., Schmidt, K. 2000. Fairness, incentives, and contractual choices. *Europ. Econ. Rev.* 44, 1057-68.

Frey, B. 1993. Does monitoring increase work effort? The rivalry between trust and loyalty. *Econ. Inquiry*. 31, 663-70.

Frey, B., Jegen, R. 2001. Motivation crowding theory. *J. Econ. Surveys*. 15, 589-611.



Frey, B., Oberholzer-Gee, F. 1997. The cost of price incentives: An empirical analysis of motivation crowding-out. *Amer. Econ. Rev.* 87, 746-55.

Gneezy, U. 2003. The W effect of incentives. Chicago. *Mimeo.*

Gneezy, U., Rustichini, A. 2000a. Pay enough or don't pay at all. *Quart. J. Econ.* 115, 791-810.

Gneezy, U., Rustichini, A. 2000b. A fine is a price. *J. Legal Stud.* 29, 1-17.

Guerra, G. 2002. Crowding out trust: The adverse effect of verification: An experiment. University of Oxford, Economics Discussion Paper, 98.

Houser, D., Xiao, E., McCabe, K., Smith, V. 2007. When punishment fails: Research on sanctions, intentions and non-cooperation. *Games and Economic Behavior*, Forthcoming.

Kachelmeyer, S., Shebata, M. 1997. Internal auditing and voluntary cooperation in firms: A cultural experiment. *The Accounting Review.* 72, 407-31.

Kreps, D. 1997. Intrinsic motivation and extrinsic incentives. *Amer. Econ. Rev.* 87, 359-64.

Laffont, J-J., Martimort, D. 2002. The theory of incentives: The Principal-Agent Model. Princeton University Press, Princeton.

Lepper, M., Greene, D. eds, 1978. The hidden cost of reward: New perspectives on psychology of human motivation. Hillsdale: Erlbaum.

Lin, C-C., Yang, C. 2006. Fine enough or don't fine at all. *J. Econ. Behav. Organ.* 59, 195-213.

Minkler, L. 2004. Shirking and motivations in firms: Survey evidence on worker attitudes. *Int. J. Ind. Organ.* 22, 863-84.

Montmarquette, C., Rulliere, J-L., Villeval, M-C., Zeiliger, R. 2005. Redesigning teams and incentives in a merger: An experiment with managers and students. *Management Sci.* 50, 1379-89.

Nagin, D., Rebitzer, J., Sanders, S., Taylor, L. 2002. Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *Amer. Econ. Rev.* 92, 850-73.

Prendergast, C. 1999. The provision of incentives in firms. *J. Econ. Lit.* 37, 7-63.

Sally, D. 1995. Conversation and cooperation in social dilemmas. A meta analysis of experiments from 1958 to 1992. *Rationality and Society.* 7, 58-92.

Schulze, G., Frank, B. 2003. Deterrence versus intrinsic motivation: Experimental evidence on the determinants of corruptibility. *Econ. Governance.* 4, 143-60.

Titmuss, R. 1970. *The Gift Relationship.* Allen and Unwin, London.

Zeiliger, R. 2000. A presentation of Regate, internet based software for experimental economics. <http://www.gate.cnrs.fr/~zeiliger/regate/RegateIntro.ppt>, GATE.

## APPENDIX A

### Instructions for the Variable Treatment (order: *Distant* relationship then *Interpersonal relationship*)<sup>20</sup>

You are going to participate in an experiment which is part of a scientific program supported by the GATE research institute of the CNRS (*National Center for Scientific Research*), and by Utah State University in the U.S.A. During this experimental session, you are requested to make decisions and you can earn money. The amount of your earnings depends on your own decisions and on those of the other participants with whom you will interact.

This session consists in 2 parts of 10 periods each. The session should last about one hour. During this session, your payoffs will be calculated in ECU (Experimental Currency Units) and put on an account.

- At the beginning of the session, your account will be credited of 150 ECUs, which are given as an initial endowment.
- During each period, you can earn or lose ECUs. Please note that your decisions can avoid losses with certainty and that possible losses in some periods should be compensated for by earnings in other periods.
- Your final earnings are equal to the sum of the ECU you will earn in each of the 20 periods. At the end of the session, the total amount of ECU you have earned on your account will be converted to Euros at the following rate:

$$150 \text{ ECU} = 1 \text{ €}$$

Your entire earnings in Euros will be paid in cash in a separate room to preserve confidentiality.

Throughout the entire session, **talking is not allowed** except when invited by the experimenter. Any violation of this rule will result in being excluded from the session and not receiving payment. If you have any questions regarding these instructions, please raise your hand. Someone will answer your questions privately.

During this session, the group of participants is subdivided into two categories in equal number: X and Y participants. Your computer indicates whether you are an X- or a Y-participant. Whether X or Y, you keep the same role throughout this session.

### Rules for Periods 1-10

During each period, pairs of participants are randomly formed (each X-participant is matched with one Y-participant). At each new period, new pairs are randomly formed. You are not necessarily matched with the same person from one period to the other. Nobody will be informed of the identity of the participants s/he interacted with during these periods.

---

<sup>20</sup> The other sets of instructions corresponding to the other treatments are available upon request to the authors.

## Roles

- The X-participant asks the Y-participant to realize a task in exchange for which s/he will receive a payment. S/He can apply a monitoring probability to Y's result.
- The Y-participant performs a task and achieves a result.

## What does occur in each period?

Each period consists of three stages:

- *First stage:* the **X-participant** decides on whether the result of the Y-participant will be submitted to an audit at the end of the period. S/He chooses this probability of the audit among the following values:

(0; .1; .2; .3; .4; .5; .6; .7; .8; .9; 1)

Examples: if s/he chooses the 0 value, this means that no audit will be done. If s/he chooses the .1 value, this means that the Y-participant's output has 1 chance out of 10 to be checked. If s/he chooses the 1 value, this means that the Y-participant's output will be checked with certainty.

Each probability is associated with a cost in ECUs that has to be borne by the X-participant, as indicated in the following Table:

Probability	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
Audit Cost paid by X	0	0	0	3	8	15	24	36	50	66	84

It should be noted that this cost is to be paid by the X-participant, should the result of the Y-participant be actually checked or not.

- *Second stage:* the **Y-participant** is informed about the choice by X of the audit probability. In each period, which lasts 1 minute, each Y-participant has to perform a task on his or her computer to achieve an outcome or result.
  - Nature of the task to be performed

This task consists in uncovering a curve where a line has been plotted beforehand. This curve is increasing and/or flat. It can have single or multiple plates that are ranked from the lowest to the highest. The highest altitude that can be reached by this curve, measured in points, is 100. The Y-participant uncovers the line of this curve as s/he moves along. Starting from point 0, s/he is making progress at the same time in terms of distance (one goes along the horizontal axis) and in terms of altitude (one goes up on the vertical axis).

The Y-participant moves by clicking one of the two buttons offered on his computer screen. These two buttons correspond to two available speeds.

- A first button enables the Y-participant to take “steps of 1”. Steps of 1 do not cost money.
- A second button enables the use of “steps of 2”. These steps are twice as rapid as steps of 1, but they cost money. Their cost is determined according to the following principle:
  - ✓ Each of the first 10 steps of 2 costs .4 ECU
  - ✓ Each of the next 10 steps of 2 costs .6 ECU
  - ✓ Each of the next steps of 2 beyond the 20th costs 1 ECU.

It is possible to switch speed at will and as many times as desired. As long as the Y-participant does not want to change speed, s/he holds the mouse down and the progression along the curve automatically proceeds at the chosen speed.

The Y-participant can stop the progression whenever s/he likes, even before the one-minute time is over. During the progression, the participant is informed of the current height reached, cumulated cost of steps of 2 and remaining time.

- The result

One names « result » the height reached when the Y-participant stops her/his progression or when the time is over; no matter the distance from the origin, only the height reached matters. This result is expressed in points and can take all integer values between 0 and 100. The maximum height of the curve being 100, the maximum result is thus also 100 points.

- *Third stage*: the computer program applies the audit probability to the result of the Y-participant and determines the participants' payoffs.

4 cases may occur.

- If the result is not audited, Y receives 100 ECUs.
- If audited and the result is between 75 and 100 points, Y receives 100 ECUs.
- If audited and the result is between 40 and 74 points, a penalty of 40 ECUs is applied: Y receives then 60 ECUs instead of 100 ECUs.
- If audited and the result is lower than 40 points, a penalty of 80 ECUs is applied: Y receives then 20 ECUs instead of 100 ECUs.

A summary table on the participants' computer screens indicates for each past period the following elements:

- The audit probability chosen by the X-participant
- The result reached by the Y-participant
- The existence of an audit when it occurred
- The payoffs for the period of the X- and Y-participants of the pair.

A new period starts automatically. Each new period is independent of the preceding periods. With each new period is associated a new curve.

### Determination of the earnings in each period

- The payoff of the X-participant is determined as follows:
  - 2.5 times the result in points reached by the Y-participant
  - the cost of the chosen audit probability
  - the amount paid to the Y-participant after deduction of the penalty, if applicable, as explained above.
- The payoff of the Y-participant is determined as follows:
  - 100 ECUs
  - the cost of the steps of 2 used by the Y-participant
  - the penalty, if applicable.

It should be noted that, **on average**, reaching the result of 75 points requires 30 steps of 2, entailing a cost of 20 ECUs. These payoffs are calculated by the computer. Participants do not have to make any calculation by themselves.

### Rules for periods 11 – 20

*[This part of the instructions was distributed only when the first part of the session has been completed].*

The task is the same as in the first 10 periods. The payoffs are determined in a similar way. As previously, a new curve appears at each new period.

The only difference is that from now on, you will interact with the same person until the end of the session.

We will indicate to you the participant with whom you will interact during the next ten periods. You have 5 minutes to introduce each other before the starting of period 11. To introduce each other, we ask you to fill a single presentation questionnaire together: the X-participant will write the answers to the questions s/he asks to the Y-participant with whom s/he is matched and the Y-participant will write the answers to the questions s/he asks to the X-participant with whom s/he is matched.

**Attention:** during this introduction, it is strictly forbidden to talk about the experiment and to conclude agreements, under penalty of exclusion from the experiment.

From the beginning of period 11 to the end of the session, you are not allowed to communicate either with your partner or with any other participant.

*[Example of the introduction sheet]*

X- Participant (to be filled in by the Y-participant) Computer Id: .....	Y- Participant (to be filled in by the X-participant) Computer Id: .....
First name :  School :  Number of brothers and sisters :  Favorite music groups:  Hobbies :  Location of vacations last Summer:	First name :  School :  Number of brothers and sisters :  Favorite music groups:  Hobbies :  Location of vacations last Summer:

**APPENDIX B (on request): *Theoretical Predictions***

**The Variable Treatment.** A risk-neutral and selfish agent performs above the minimum effort requirement as long as the expected benefit outweighs the expected costs. Recall that wage levels are discrete, and so an agent who chooses to shirk will either produce  $\underline{y}$  or 0. We have two no-shirking conditions depending on the reference level of output,  $\hat{y}$  or  $\underline{y}$ :

No-shirking condition 1 (agent  $i$ ):

$$\pi_i(y_i \geq \hat{y}) = W - c(\hat{y}) \geq E\pi_i(y_i < \underline{y}) = m[w - c(y_i)] + (1-m)[W - c(y_i)] \quad (4)$$

No-shirking condition 2 (agent  $i$ ):

$$\pi_i(y_i \geq \hat{y}) = W - c(\hat{y}) \geq E\pi_i(\underline{y} \leq y_i < \hat{y}) = m[w - c(y_i)] + (1-m)[W - c(y_i)] \quad (5)$$

Since  $c(\underline{y}) = c(0) = 0$ , the binding constraint for reaching  $\hat{y}$  simplifies as:

$$c(\hat{y}) \leq m(W - w) \quad (6)$$

That is, the marginal cost of  $\hat{y}$  must be less than or equal to its marginal benefit.

In addition, a rational and selfish agent should never perform more than  $\hat{y}$  with costly steps since it yields no additional earnings to the agent. The best reply output choices  $y_i^*$  to each monitoring intensity depends on the comparison between expected marginal cost and marginal benefit of choosing  $\hat{y}$  vs.  $\underline{y}$ . With our parameterization and an average cost of 20 to reach the desired output level, the best reply effort changes at  $m = .5$ . For  $m < .05$ , the marginal cost of  $\hat{y}$  exceeds its marginal benefit so that the agent should choose  $\underline{y}$  for  $0 < m < .5$  (or  $y=0$  for  $m=0$ ). The optimal output in the Variable Treatment is:

$$y_{iV}^* = \begin{cases} 0 & \text{if } m = 0 \\ \underline{y} & \text{if } m < .5 \\ \hat{y} & \text{if } m \geq .5 \end{cases} \quad (7)$$

An incentive compatible contract leads to  $y_{iV}^* = \hat{y}$  for monitoring  $m \geq .5$ .

Consider now the decision of a profit-maximizing principal. The principal chooses the monitoring intensity that maximizes her expected payoff considering the rational choice of the agent. This maximization can be written as:

$$\underset{m_V}{\text{Max}}(E\pi_{PV}) = m_V [v\tilde{y}_i - c(m_V) - \tilde{w}] + (1-m_V)[v\tilde{y}_i - c(m_V) - W] \quad (8)$$

$\tilde{y}_i$  is the expected output from agent  $i$  and  $\tilde{w}$  is the expected wage depending on the verified level of output. With probability  $m_V$ , output is audited and  $W$ ,  $w$  or  $\underline{w}$  is paid



depending on actual output produced. With probability  $(1 - m_V)$ , output is not verified and  $W$  is paid. Monitoring is costly whether or not shirking is detected.

In equilibrium,  $m_V$  should be chosen such that the marginal cost of monitoring equals its expected marginal return in terms of both penalty (i.e., extra profit) if shirking is detected and of additional output (increased monitoring may stimulate effort). Since both the output and monitoring are discrete choices, we can again only consider the numerical solution of the game. With our parameterization, the principal reaches her maximum expected payoff when choosing  $m_V = .5$ . Thus, in equilibrium, one obtains:

$$\begin{aligned} m_V^* &= .5 \\ y_{iV}^* &= \hat{y} = 75 \\ \pi_{iV} &= 100 - 20 = 80 \\ \pi_{pV} &= 187 - 15 - 100 = 72 \end{aligned}$$

**The Fixed Treatment.** Regarding the agent's decision, the same no-shirking conditions apply as in the other treatment:  $c(\hat{y}) \leq m(W - \underline{w})$ . Optimal output is thus the same as in equation (9):

$$y_{iF}^* = y_{iV}^* \tag{9}$$

The principal chooses the monitoring intensity that maximizes her expected payoff considering the rational choice of the agent. This maximization can be written as:

$$\underset{m_F}{Max}(E\pi_{pF}) = m_F [f - c(m_F) - \tilde{w}] + (1 - m_F) [f - c(m_F) - W] \tag{10}$$

with  $\tilde{w}$  the expected wage depending on the verified level of output. With our parameterization, the principal reaches her maximum expected payoff when choosing  $m_F = .3$ . Thus, in equilibrium, one obtains:

$$\begin{aligned} y_{iF}^* &= \underline{y} = 40 \\ m_F^* &= .3 \\ E\pi_{iF} &= (0.3) * 60 + (0.7) * 100 = 88 \\ E\pi_{pF} &= (0.3) * (180 - 3 - 60) + (0.7) * (180 - 3 - 100) = 89 \end{aligned}$$

A selfish principal chooses a low probability of monitoring (.3) in the *Fixed* treatment. This contract is not efficient since the agent should stop his effort as soon as he reaches  $\underline{y}$ , which he can always reach without bearing any cost. Compared to the *Variable* treatment, we should thus observe a lower monitoring intensity and a lower output.