A 'threat' is a 'Threat': Incentive Effects of Firing Threats with Varying Degrees of Performance Information

Jordi Brandts
Brice Corgnet
Roberto Hernán-González
José M. Ortiz
Carles Solà

February 2018

Barcelona GSE Working Paper Series

Working Paper nº 1023

A ‘threat’ is a ‘Threat’:
Incentive Effects of Firing Threats with Varying Degrees of Performance Information

Jordi Brandts (IAE(CSIC) and Barcelona GSE)
Brice Corgnet (EM Lyon, GATE)
Roberto Hernán-González (Burgundy School of Business)
José Mª Ortiz (Middlesex University)
Carles Solà (Universitat Autònoma de Barcelona)

February 2018

Abstract

We study the incentive effect of firing threats when bosses have limited information about workers. We show that a minimal amount of individual information about workers’ effort such as the time spent at their work station is sufficient to ensure strong incentive effects. This supports the use of firing threats based on rudimentary yet uncontroversial measures of work performance such as absenteeism, in organizational settings in which only limited information about workers is available. Our results help understand the limited link between pay and performance observed in compensation contracts calling for an extension of the principal-agent model to take into account how workers (mis-)perceive the intensity of incentives.

JEL classification: C92, D23, D82
Keywords: Firing Threats, Incentives, Informativeness Principle, Laboratory xxperiments

Acknowledgements: We thank the Spanish Ministry of Economics and Competitiveness through Grant: ECO2017-88130 and through the Severo Ochoa Program for Centers of Excellence in R&D (SEV2015-0563) and the *Generalitat de Catalunya* (Grant: 2017 SGR 1136) for financial support.

1. Introduction

In the economic theory of incentives precise information about workers' individual effort is regarded as a key input for the design of efficient compensation contracts (e.g. Laffont and Martimort, 2002; Bolton and Dewatripont, 2004). In the absence of such precise individual measures of output, firing threats have been proposed as an essential feature of the optimal employment contract (Becker and Stigler, 1974; Klein and Leffler, 1981; Shapiro and Stiglitz, 1984; MacLeod and Malcomson, 1989). To assess the effectiveness of firing threats as an incentive mechanism one thus has to consider a setting in which such threats are likely to be used. Such setting is one in which supervisors possess limited information about workers' effort (e.g. Shapiro and Stiglitz, 1984). In this paper we study behavior in such an environment. By assessing the effectiveness of firing threats across different levels of information, our work also aims at assessing the 'robustness' of incentive effects to a wide range of circumstances following the early call of Hart and Holmström (1986).

To conduct our study, we rely on laboratory experiments so as to precisely vary the quantity of information across the distinct treatments. The experimental methodology allows us to assess the causal link between the quantity of information available to supervisors and the effectiveness of firing threats. To our knowledge, such a test has never been performed. Related to our work is the literature studying the effect of different incentive schemes on costly monetary transfers. For example, Nalbantian and Schotter (1997) studies monitoring incentive schemes in which the monetary transfer of the agent is observed with a given probability. If the level of transfer is observed to be below a preset level, the agent is given a low pay. Nalbantian and Schotter (1997) show that increasing the intensity of monitoring, by increasing the probability a transfer is observed, generates a large positive effect on transfers. In Nalbantian and Schotter (1997), there was no principal and the monitoring technology was exogenously defined.

In the organizational context that we study in this paper there is a boss in charge of both actively monitoring and paying workers who can choose between three activities: doing a real-effort task, spending their time chatting or browsing the web. Our setting uses a lab workplace à la Corgnet, Hernán-González and Schniter (2015). In order to isolate the incentive effect of firing threats, we consider dismissals as the only incentive mechanism available to bosses. Workers receive a fixed wage each period which can not be made contingent on the information available to the boss.

We conducted four treatments, three of which allow the boss to fire workers. In the firing treatments, the boss could fire one (out of nine) workers at the end of each of five periods (except for the first period). Firing treatments differed in the amount of information which could be collected by the boss during the monitoring of workers.

In the *complete information* treatment, bosses had access to real-time information about each worker's production as well as about the current activity workers were undertaking (either working, chatting or browsing the internet). In the *partial information* treatment, bosses could not observe workers' production levels but could see the current activity each of them was undertaking. Finally, in the *minimal information* treatment, bosses could neither observe workers' production nor could they observe their current activity. They were only informed about the total production of the organization. In the baseline treatment, the boss had complete information but could not fire anyone.

We formulate conjectures based on a multi-period principal-agent model in which some agents have an intrinsic motivation for work and the principal can fire agents at the end of each period. In line with the *informativeness* principle (Holmström, 1979), our model predicts that the production of the organization should be highest under *complete information*. In addition, *minimal information* should lead to levels of effort which are not substantially higher than in the *baseline* without firing.

We find that workers' production and task dedication are significantly higher in all three firing treatments than in the b*aseline* in which firing workers is not possible. In particular, organizational production is more than 50% higher in the *minimal information* treatment compared to the *baseline*. Organizational production in the *minimal information* treatment is, however, about 20% lower than in the other two firing treatments. *Partial* and *complete information* lead to the same organizational production. Our results are not consistent with the predictions of the principal-agent model we use, since firing threats are effective even when there is no individual information about workers' actions (as is the *minimal information* treatment) and additional information is valuable, without the need for it to be complete.

We also find no evidence for the relevance of monitoring and control costs, social incentives and intrinsic motivation. This brings us to the title of paper, 'a threat is a Threat', which is meant to convey that it is purely the salience of the firing threat and workers' reaction to it which explains what we find.

Our results can help understand the lack of empirical support for the *informativeness* principle (e.g. Baker, Jensen and Murphy, 1988; Chiappori and Salanié, 2000) which Holmström himself acknowledges (Hart and Holmström, 1986, page 51). Traditional explanations of the weak link between effort and pay in firms stress the negative effects of excessive monetary incentives which can negatively affect intrinsic motivation (e.g. Bénabou and Tirole, 2003), prevent workers from engaging in multitasking (Holmström and Milgrom, 1991) or exploring new ideas (e.g. Manso, 2011).[1] Our findings put forward a novel explanation for the *informativeness* principle puzzle which relies on the fact that a tighter link between effort and pay is not necessary for compensation schemes to provide strong incentives. Our results suggest that firing threats may be popular within organizations because their effects are robust to different informational conditions. These findings are also consistent with the theoretical literature putting forward scarce information as a necessary condition for firing threats to be optimal (e.g. Shapiro and Stiglitz, 1984).[2]

2. Conceptual framework

In what follows we present a simple model that serves as a theoretical basis for our main conjecture. We consider a setting in which one principal pays each of the n agents a fixed wage (w) in each of T periods. The principal also has the possibility to fire a proportion f of the agents at the end of each period. Each worker i decides to allocate his or her time, normalized to one, to the following activities: productive effort ($e_i^P \in \{0,1\}$), non-productive effort ($e_i^{NP} \in \{0,1\}$) or leisure ($l_i \in \{0,1\}$), so that $e_i^P + e_i^{NP} + l_i = 1$.

[1] Explanations for the limited use of pay-for-performance contracts have built on the fact that workers may hold social motives such as altruism (e.g. Rotemberg, 1994, Dur and Sol, 2010) or inequity aversion (Bartling and von Siemens, 2010). Also, workers who are intrinsically motivated may react negatively to powerful incentives (e.g. Bénabou and Tirole, 2003) or may be motivated by alternative non-monetary compensation such as symbolic rewards (e.g. Kosfeld and Neckermann, 2011), status (e.g. Charness et al. 2014), the mission of the firm (e.g. Besley and Ghatak, 2005) or non-binding goals (e.g. Wu et al. 2008; Goerg and Kube, 2012; Corgnet et al. 2016). Peer pressure may ensure high levels of effort without needing high-powered incentives (e.g. Kandel and Lazear, 1992). When accurate performance measures are only available for a subset of worker's relevant tasks, managers may weaken incentives to ensure that workers will not only dedicate their effort to complete the task for which performance measures are available (Holmström and Milgrom, 1991). Weak incentives may also be optimal when the organization wants workers to explore new alternatives (e.g. Manso, 2011). Relational contracts may also help sustain optimal low-powered incentive schemes (e.g. Levin, 2003). Finally, low-powered incentives may be optimal to avoid disagreements between managers and workers in a context in which they have different prior beliefs regarding the firm's available projects (van den Steen, 2010).

[2] See Falk, Huffman and MacLeod (2015) or Charness, Cobo-Reyes, Jimenez, Lacomba and Lagos (2017) for experimental literature studying dismissal barriers.

Agents' productive effort determines their individual output ($q_i := e_i^P$) where $i \in \{1, \dots, n\}$. The total output produced by agents is defined as $y = \sum_{i=1}^{n} q_i$. Non-productive effort simply consists in being present at the work station without working and thus entails no production and no effort costs. The utility function of agent *i* in a given period can be expressed as follows:

$$U_i \coloneqq w - C(e_i^P) + \eta(l_i)$$

where $C(e_i^P)$ stands for the cost of productive effort function with $C' > 0$ and $C'' > 0$, $\eta(l_i)$ stands for the utility of leisure with $\eta' > 0$ and $\eta'' < 0$. When fired, the utility of the agent becomes $\eta(l_i)$. We derive our main prediction using the following specification of the agents' utility function:[3]

$$U_i \coloneqq w - \alpha_i \frac{{e_i^P}^2}{2} - \beta_i \frac{(1 - l_i)^2}{2}$$

We consider two types of agents which can be described as intrinsically motivated (α_i^I <0) or not ($\alpha_i^{NI} > 0$). The proportion of intrinsically-motivated agents in the population is commonly known to be π_I. All agents enjoy the leisure alternative in the same direction β_i>0.

The principal maximizes the difference between total organizational output and the sum of wages paid to agents. To that end, the principal can fire a proportion *f* of agents at the end of each period We consider three cases, one in which the principal only observes total production (y) (*minimal information*), one in which information about workers' leisure activities (l_i) (*partial information*) is also available and one in which the principal can observe workers' individual production levels (q_i) along with total production and leisure activities (*full information*).

In the baseline case in which the principal cannot fire (*f* = *0*), only intrinsically-motivated agents will exert productive effort. Our aim is thus to study the extent to which firing threats can induce non-intrinsically motivated agents to exert productive effort. These conditions will depend on the amount of information available to the principal. In line with the *informativeness* principle, non-intrinsically motivated agents will exert productive effort for a wider range of parameters when more information is available to the principal (see Appendix O.2 for details).

[3] For ease of exposition, we express the utility of leisure (Internet browsing) as the opportunity cost of not browsing the Internet $(1 - l_i)$.

The model yields conjectures which we can check in the following controlled setting that mimics the essential features of our model. More broadly, the informativeness principle suggests that every piece of information can be valuable and this suggests that the more information the higher the production of the organization will be.

3. Design

3.1. *Lab workplace*

Our computerized experimental environment consists of a virtual organization with one boss (referred to as C) and nine employees (referred to as Bs).[4] In our setting, employees can, at any point in time during the experiment, complete a real-effort task, access the Internet for leisure purposes or chat with other employees. Each of the three activities is undertaken in a separate window so that only one activity can be completed at a time allowing the experimenter a precise measurement of the time spent on each activity by each subject.

The software allows for the boss to monitor employees' activities in real time and to track their experimental IDs across periods. A session consisted of five production periods of 20 minutes each. The length of the experiment was chosen so as to be able to observe fatigue and uncover incentive effects (e.g. Corgnet, Hernán-González and Schniter, 2015).

3.1.1. *The work task*

We use a particularly long and laborious *work task* to ensure that working on it required a significant level of effort. All subjects, the employees and the boss, had to add up numbers from tables with 36 numbers for one hour and 40 minutes.[5] The reason for having the boss work on the same task as the employees is to allow him or her to assess its difficulty and, thus, to make firing decisions knowingly.

In the *work task*, subjects were not allowed to use a pen, scratch paper or calculator. This rule amplified the level of effort subjects had to exert in order to add up the matrices correctly. Each table had six rows and six columns. The numbers in each table were generated randomly.

---

[4] We chose to have ten people in each organization so as to represent a small company, which both in the EU and the US consists of at least 10 people.

[5] Different variations of this task have been used by Bartling et al. (2009), Dohmen and Falk (2010), and Abeler et al. (2011). A counting task that consisted of summing up the number of zeros in a table randomly filled with ones and zeros was also used in Falk and Huffman (2007). A long typing task was used in Dickinson's (1999) experiment for which subjects had to come during four days for a two-hour experiment. Falk and Ichino (2006) used a four-hour mailing task in their field experiment on peer effects. In another field experiment by Gneezy and List (2006), subjects were asked to enter data into a computer database for six hours.

Each table completed correctly generated a 40-cent profit while a penalty of 20 cents was subtracted from individual production for each incorrect answer.[6] At the end of each period, the total amount of money generated by all ten subjects during the period was displayed in the history panel located at the bottom of their screens.

3.1.2. *Internet browsing*

The Internet browser was embedded in the software so that the experimenter could keep a record of the exact timing of the activities completed by each subject. Subjects were informed that their usage of the Internet was strictly confidential.[7]

The introduction of the possibility of using the Internet is motivated by the widespread use of Internet at the workplace (Malachowski, 2005). An appealing feature of Internet as an alternative to the *work task* is the wide range of activities that can be undertaken online. Many people are likely to derive utility from Internet access as they will be able to browse web pages that best fit their personal interests.[8]

3.1.3. *Chatting activities*

The boss and all employees also had access to a chat room through which they could communicate with the other subjects during the experiment. A subject could send a message to all subjects at once or to any subset of them. If subjects received a message while not currently in the chat room, a pop-up window displaying the content of the message as well as the experiment *ID* of the sender would automatically appear on their screen. A subject would, however, have to enter the chat room to send a message. As a result, incoming chat could potentially distract subjects completing the *work task*.

3.1.4. *Monitoring activities*

In all firing treatments but the one with *minimal information*, the boss could monitor the nine employees' activities at any time during the experiment. Monitoring activities had to be undertaken in a separate window so that bosses could not complete their own *work task*, *chat* or *browse the Internet* while monitoring their employees. When on the monitoring screen, bosses could decide whether to monitor all or a subset of the employees at the same time. Depending on

[6] Penalties did not apply when individual accumulated production was equal to zero so that individual production could not be negative.
[7] Subjects were expected to follow the norms set by the university regarding the use of Internet on campus.
[8] Two related studies (Eriksson, Poulsen and Villeval, 2009, Charness, Masclet and Villeval, 2014) have also introduced on-the-job leisure activities in experimental environments by giving subjects access to magazines.

the treatment, the monitor received information in real time about the activities undertaken by the selected subject, their current total production, as well as their contribution to the *work task* (in % terms). This virtual monitoring activity was designed to mimic current organizational technology (e.g. SpectorSoft 360, Virtual Monitoring™, Employee Monitoring) that allows for real-time monitoring of employees' activities by tracking the time they spend on various applications.

Whenever they were being watched, employees were notified with a message stating "*The C subject is watching you*" jointly with an eye picture. At the end of each period, the boss had access to a monitoring summary which, depending on the treatment, included information regarding employees' activities during the period, their production levels as well as their contribution to total production.

In addition to the previously mentioned activities, each subject could click on a box moving slowly from left to right at the bottom of their screen. Each time subjects clicked on a box they earned *5 cents*. The box appeared at the bottom of a subject's screen every 25 seconds independently of whether the subject was currently working on the work task, chatting, or browsing the Internet. Given that the experiment consisted of 5 periods of 20 minutes each, subjects could earn a total of $12.00 just by clicking on all the 240 boxes that appeared on the screen during the experiment. This aimed at representing the pay that workers obtain just for being present at their workstation regardless of their commitment to the *work task*. The rationale for this task was to create an environment in which subjects perceive that showing up for work already paid off. This was introduced to help attenuate *active participation*, an issue which was raised by Lei, Noussair and Plott (2001) in the context of experimental asset markets. In the context of our lab workplace, Corgnet, Hernán-González and Schniter (2015, page 285) stress that "*subjects may engage actively in a focal work task because of expectations, rewards, and lack of desirable alternatives*". When desirable alternatives are present, active participation in effortful work may be traded off to some degree, revealing subtle incentive effects.

3.2. *Treatments*

Table 1 shows the main features of our treatments together with the number of subjects in each cell. In all treatments, employees were rewarded a fixed wage of 200¢ each period; they were not incentivized based on their performance on the *work task*.[9] The boss received the output produced by all subjects (including himself) on the *work task*, but was not paid a fixed wage. In all three firing treatments (*complete, partial* and *minimal info*), the boss could fire one employee at the end of each of periods 2, 3 and 4.[10] The boss kept the fixed wage of dismissed employees in the following periods. Our aim was to conduct a conservative test for the effects of firing threats on employees' production by considering that only one of them could be fired each period. Hence, up to three out of nine employees could be fired in a given experiment.

Table 1.
Summary of the treatments.

Treatment	Description	Number of sessions (subjects)
No Firing Complete info (*Baseline*)	Employees were paid a fixed wage of 200¢ per period. The boss subject kept the value of all output produced by all employees in the organization. In addition, Bosses were paid the value of their own production. Bosses could monitor employees' activities and individual production but had no possible recourse.	6 (60)
Firing *Complete info*	The boss could monitor employees' activities and individual production, and could fire one employee at the end of periods 2, 3 and 4. Payment as in in the baseline but the boss also kept the fixed wage of dismissed employees.	6 (60)
Firing *Partial info*	Same as *complete info* except that bosses could *only* monitor employees' activities not accessing any information regarding their individual production.	6 (60)
Firing *Minimal info*	Same as previous firing treatments except that bosses could *not* monitor employees and thus only had access to the total production of the organization when deciding upon firing employees.	6 (60)

[9] The choice of 200¢ was made so that, at least some *employees* would not be able to produce that value thus inducing the boss to fire workers. This value was calibrated using previous related experiments (e.g. Corgnet, Hernán-González and Rassenti, 2015a).
[10] We do not allow for firing after period 1 because of the large learning effects observed in the great majority of real-effort experiments that makes the first period substantially different from the rest of the experiment (e.g. e.g. see Charness and Campbell, 1988).

Dismissed employees could only browse the Internet. They were rewarded solely for their earnings on the clicking task which were reduced to 1¢ per box instead of 5¢ per box for the *active* employees and the boss.[11] They were not able to chat with the members of the organization, and they could not be rehired.

3.3. *Survey data*

For each session, we collected survey data for a number of items (see Appendix O.1). This information was used to provide controls for our statistical analysis and to investigate our findings further.

Adding skills. Subjects were asked to sum as many sets of five one-digit numbers as they could during two minutes in the spirit of Dohmen and Falk (2011). Each correct answer was rewarded 10 cents. The number of correct answers is what we refer to as "ability". To ensure that this measure was not affected by fatigue and treatment differences, it was collected upon arrival at the lab and before receiving instructions for the corresponding treatment.

At the end of the experiment, subjects were asked to fill out a 10-minute survey including questions regarding demographics, cognitive skills and social preferences. We collected these measures at the end of the experiment because they are less central to our study than adding skills and were not planned to be used as main controls in our analysis.

Demographics. We asked subjects about their name, age and gender. We also asked them how many hours a week they usually worked for pay or volunteer. We also collected data regarding which degree they were currently studying.

Cognitive skills. We measured cognitive reflection using the CRT developed by Frederick (2005). Our CRT measure sums the number of correct answers on the test.

Social preferences. We elicited social preferences following Bartling et al. (2009). We asked subjects to make six choices between two possible allocations of money between themselves and another anonymous and randomly assigned subject in the experiment. In each experimental session, two subjects and one of the six decisions were selected at random for payment. The choice of the first subject in the selected decision was used to allocate payoffs between the two

[11] As a result, the maximum period earnings of dismissed subjects on the clicking task were equal to 48¢ instead of 240¢ for *active* employees and the boss.

subjects. All decisions were anonymous. The allocation decisions are described in Table O1 in the online appendix.

Intrinsic motivation. We measured subjects' intrinsic motivation at the end of the organizational experiment by assessing whether subjects were willing to sum sets of five one-digit numbers for two minutes in the absence of any monetary rewards (Dohmen and Falk, 2011). This measure was purposefully collected after the treatment was completed as it was aimed at assessing the effect of the treatment on workers' motivation to keep producing for no monetary incentives. We measure intrinsic motivation as the ratio of their performance on this task and their performance on the incentivized version of the same two-minute task completed when they entered the lab.

3.4. *Procedures*

Our subject pool consisted of students from two major Spanish Universities. The experiments took place between December 2014 and June 2016. In total, 240 subjects participated in the experiments, divided into 24 groups of 10 subjects each, that is six groups for each treatment. All of the interaction was anonymous. Subjects had 20 minutes to read the instructions on their screens. Three minutes before the end of the instructions period, a monitor announced the time remaining and handed out a printed copy of the summary of the instructions. None of the subjects asked for extra time to read the instructions. The interaction between the experimenter and the subjects was negligible.

At the end of the experiment, subjects were paid their earnings in cash, rounded up to the nearest quarter. Individual earnings at the end of the experiment were computed as the sum of all earnings in the *5* periods plus the earnings from the adding and social preferences tasks included in the questionnaire. Participants playing the role of an employee (boss) in the *complete, partial*, *minimal info* and *baseline* treatments earned €29.36 (€97.47), €28.06 (€95.58), €29.21 (€76.54), €29.09 (€54.26) on average, respectively. This includes a five euro show-up fee. Experimental sessions lasted on average two hours and thirty minutes.

4. Results

In sections 4.1 and 4.2 we analyze the first four of the five periods of our experiment which correspond to the periods in which firing threats had monetary consequences. In Section 4.4.3, we present the results of the last period to assess the motivational implications of firing threats. In Appendix A we report results when pooling data for all five periods (see Tables A.7 and A.8).

4.1. *Employees' production and work dedication*

We define individual production as the monetary amount generated per answer by a given subject on the task divided by the reward for a correct answer (40 cents). At every step of the analysis, we only include subjects who belong to the organization excluding fired subjects. This means that in the three treatments where firing was possible, subjects who had been fired before an actual period were excluded from the analysis. Note, however, that none of the results are qualitatively affected by including the fired subjects in the analysis and considering their production to be zero.

Fig.1 and Fig.2 show average production and working time for the first four periods of our four treatments, respectively. Working time is defined as the percentage of their time employees spent on the task screen instead of browsing the web or chatting with other subjects. Working time is thus a measure of work dedication which negatively correlates with on-the-job leisure which can be calculated as the time spent browsing and chatting. We analyze working time as well as production as these two measures can lead to different results if employees exert non-productive effort (being present at the workstation without completing the task).

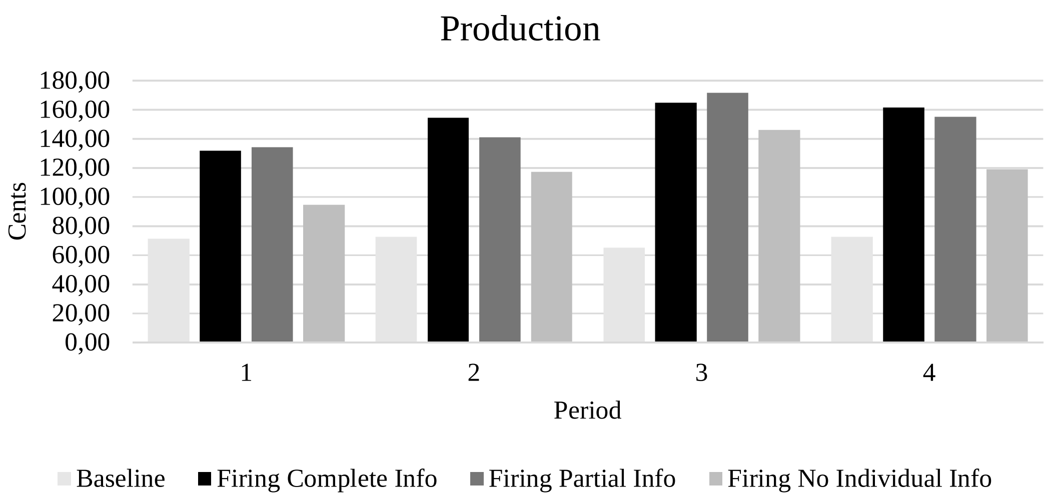

Fig.1 Employees' average production across treatments for periods 1 to 4. Subjects who have been fired before a current period are excluded.

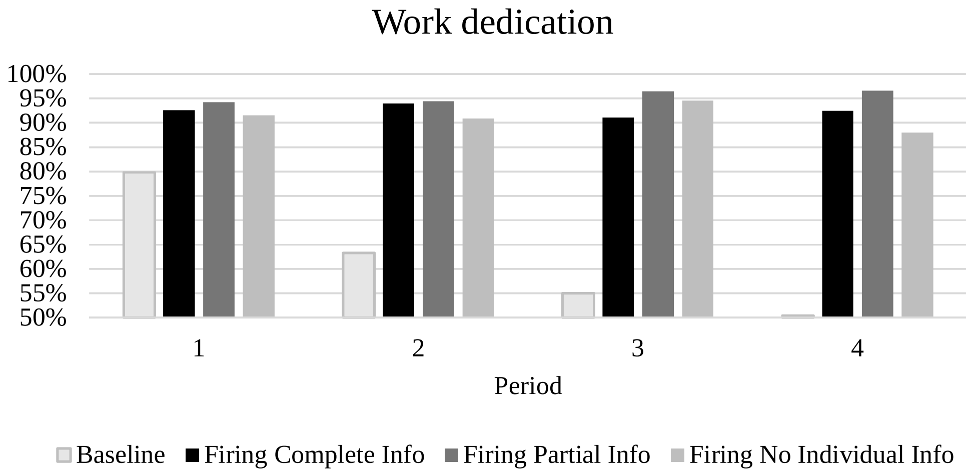

Fig.2 Employees' average working time (%) across treatments for Periods 1 to 4. Subjects who have been fired before an actual period are excluded.

We use a GLS random-effects model (see Table A.1 in Appendix A)[12] to check for any statistically significant differences in employees' individual production across treatments. Given the nature of our data, we use panel regressions and robust standard errors clustered at the session level.[13] We use random effects in all the regressions reported in this paper. Using the Breusch-Pagan Lagrange Multiplier test, we cannot reject the random effects specification

The results show that individual production for employees is significantly higher in treatments where firing is available compared with the *baseline* without firing threats. Average production in the *complete info* (152.8) and *partial info* treatments (150.4) do not differ significantly. But, the average production of employees is about 30% higher in the firing treatments in which the boss has access to individual information about employees (*complete info* and p*artial info* treatments) compared to the *minimal info* treatment. These differences are statistically significant (see Table A.1).

At the same time, employees spend significantly less time on the work task in the *baseline* (62.13%) than in the firing treatments (93.15%). In line with the production results, we do not find any significant differences in the time spent on the task between the firing treatments with

[12] Appendix B contains several tables with a robustness analysis of our regressions.
[13] Following Cameron and Miller (2011), we also estimated standard errors using the wild bootstrap procedure. Using this procedure, we obtained very similar p-values to the ones reported in the results section. In particular, the effects which are shown to be statistically significant using robust standard errors continue to be significant when using the wild bootstrap procedure.

complete info (92.51%) and *partial info* (95.38%) (see Table A.2 in Appendix A).[14] Also, consistently with the production findings, the time spent on the work task was increased in the *complete info* and *partial info* treatments compared to the *minimal info* treatment (91.25%). The differences in working time were, however, of limited magnitude and only significant when comparing *partial info* and *minimal info* firing treatments (see Table A.2).

Our findings on production and work dedication contrast with our conjecture that the incentive effect of firing threats becomes weaker when individual information about workers is less precise.

Finally, subjects could also obtain earnings from clicking on yellow boxes appearing every 25 seconds at the bottom of their screen. No significant differences were observed across treatments regarding the clicking task. Subjects successfully clicked on the box in 96.46%, 93.77%, 92.59% and 94.25% of cases in treatments *complete info*, *partial info*, *minimal info* and *baseline*, respectively.

4.2. *Firing decisions*

In this section, we analyze bosses' firing decisions in the three firing treatments. In Table 2 we show that the fired employees in a given period in the *complete* and *partial info* treatments were producing less than the rest of employees in the organization. Not surprisingly, the difference in production between fired and non-fired subjects was more pronounced for the *complete info* than for the *partial info* treatment (as indicated by the higher p-values reported in the last row of each panel). However, as expected, we do not observe production differences in the *minimal information* treatment when comparing fired employees with the rest of employees in the organization as no individual information was available to bosses.

It is worth noting that only a few employees were actually fired (less than one third of the number of employees that could have been fired). Across all the sessions of the three firing treatments, only 17 employees were fired out of 162. This is in line with our results above that show that firing threats induce high effort which in turn implies that the principal does not need to fire workers.

[14] We do not present treatment comparisons for bosses in the main text because of their limited number per treatment. However, we show these results in Tables A.7 and A.8 in Appendix A.

Table 2.

Firing decisions per period across treatments.

		Period 2	Period 3	Period 4	Total
Complete info	Total [maximum possible] number of fired subjects	2 [6]	2 [6]	3 [6]	7 [18]
	Average production of employees before being fired*	20	40	66.8	271.6
	Average production of other employees	197.6	204.4	169.2	715.2
	p-value†	2.119 (**0.0341**)	2.078 (**0.0377**)	1.744 (**0.0812**)	3.154 (**0.0016**)
Partial info	Total [maximum possible] number of fired subjects	2 [6]	3 [6]	1 [6]	6 [18]
	Average production of employees before being fired	0	100	0	156.8
	Average production of other employees	148.8	177.2	154.4	694.4
	p-value	2.12 (**0.0339**)	1.17 (0.2399)	1.537 (0.1244)	2.058 (**0.0396**)
Minimal info	Total [maximum possible] number of fired subjects	1 [6]	1 [6]	2 [6]	4 [18]
	Average production of employees before being fired	80	120	170	364.8
	Average production of other employees	114.8	160	130	520
	p-value	0.394 (0.6936)	0.659 (0.5192)	-0.479 (0.6318)	1.169 (0.2422)

†This p-value refers to the Wilcoxon rank-sum test that assesses whether average production is the same for subjects who were fired and for those who were not fired..

Our results suggest that firing threats are 'robust' incentive schemes (Hart and Holmström, 1986) which, regardless of the information available to the boss, effectively increase workers' production compared to *baseline*. It seems that firing threats always appear as salient threats regardless of the information on which they are based.

4.3. *A 'threat' is a 'Threat'*

In this section we study 'mechanisms' by which firing threats work by looking at whether they have different effects across ability levels and across periods.

4.3.1. *Firing threats across ability levels*

In the *complete info* treatment, the magnitude of the firing threats differs conceptually between low- and high- ability workers. Indeed, because a maximum of one third of the workers could be fired in an organization, high-ability workers could ensure their survival in the organization by producing slightly more than a third of the workers. We investigate this issue by classifying employees as either high or low ability subjects depending on whether their score on the adding task (used to measure their ability prior to starting the experiment) was either above or below the median performance of the subjects participating in the current study.[15] Across all treatments, 51.11% of subjects are classified as high-ability subjects and 48.75% are classified as intrinsically motivated. We do not find significant differences in the proportion of high-ability subjects across treatments (p-values > 0.10 for all pairwise comparisons).

We find that the incentive effect of firing threats holds regardless of ability levels (see Fig. C.1 in Appendix C). Both low- and high- ability workers produce significantly more in any of the firing treatments compared to the *baseline*. That is, high-ability employees reacted similarly to low-ability workers to firing threats despite that only a maximum of one third of the employees could ever be fired. In the *complete info* treatment, high-ability workers could have produced slightly more than low-ability workers to ensure their survival in the organization. However, they produced on average almost twice as much. These findings are consistent with the fact that firing threats are salient regardless of the magnitude of the threat. That is, 'a threat is a Threat'.

[15] Our results are robust to categorizing subjects' ability with respect to the median performance of a given experimental session instead of the pool of subjects recruited for the study.

4.3.2. *Firing threats across periods*

In the *complete info* treatment the magnitude of the firing threats does not only differ across ability levels, but it also differs across periods. Being fired in the second period entails larger costs in terms of foregone wages than being fired at the end of the fourth period. However, as is shown in Table A.7, workers' production is almost identical in periods 2 and 4 across all treatments. It follows that the positive effect of firing threats is as pronounced in period 2 as in period 4, regardless of the firing treatment. These findings are again consistent with the fact that firing threats are salient regardless of the magnitude of the threat. Again, 'a threat is a Threat'.

4.4. *Other explanations*

Up to this point we have documented (1) that firing threats have an effect even with minimal information, (2) that some additional information about workers' effort leads to increased performance and (3) that complete information does not yield higher performance that partial information. All this is consistent with the explanation that it is purely the salience of the firing threat which is at work. In this section we look at other mechanisms.

4.4.1. *Monitoring and control costs*

One possible explanation for the finding that complete information does not lead to higher performance than partial information is that the *complete info* treatment may have generated an excessive amount of monitoring that was detrimental to organizational production as employees could have perceived this intense supervision as distrust (e.g. Dickinson and Villeval, 2008). However, we have several pieces of evidence that do not confirm this explanation. First, the fact that bosses fired employees according to their relative performance levels in both the *complete* and *partial info* treatments (see Table 2) suggests that monitoring employees may have been as intensive in the *partial info* treatment as in the *complete info* treatment. Second, bosses spent about the same time monitoring in the *complete info* firing treatment (14.58%) as in the treatment with *partial info* (10.40%) (see Appendix A, Table A.9). This difference was not statistically significant (see Table A.10). In addition, this reasonable amount of monitoring does not seem to correspond to a case of high monitoring intensity as is described by Dickson and Villeval (2008) and which can entail distrust. Also, the negative effect of monitoring identified by Dickson and Villeval (2008) only appeared when workers had friendship ties, which is not the case in our experiments.

Finally, we show that the monitoring activities of the boss did not have a negative effect on subsequent workers' production. In Appendix D, we present additional analysis that shows that a worker who was being watched by the boss in a given time span of five minutes did not significantly modify his or her own production in the next time span of five minutes in both the *partial info* (p-value = 0.366 for the dummy variable that takes the value 1 if a worker has been watched in the previous 5 minutes) and *complete info* (p-value=0.934) treatments (see Table D.1 in Appendix D). In sum, our findings do not seem to be consistent with an explanation based on excessive monitoring and control.

4.4.2. *Social incentives: Team identity and fairness concerns*

The presence of social incentives is a potential explanation for why performance is already high with minimal information and also for why under partial information it is as high as under complete information. As is illustrated in our conceptual framework (see Appendix O.2), one notable difference between our firing treatments is that firing decisions and thus workers' compensation are expected to be a function of total production in the *minimal info* and *partial info* treatments whereas they only depend on individual production in the *complete info* treatment.[16] It follows that one's pay in the *minimal* and *partial info* treatments depend on others' effort which is an essential aspect of *social incentives* settings (e.g. Kandel and Lazear, 1992; Bandiera, Barankay and Rasul, 2010). By contrast, *social incentives* are unlikely to be present in the *complete info* treatment as pay should only depend on one's own performance. In the presence of *social incentives*, employees may exert high effort because they care about other organizational members (e.g. Rotemberg, 1994; Dur and Sol, 2010). Rotemberg (1994) stresses that altruism can help alleviate the free riding in teams problem whereas Dur and Sol (2010) stress that social incentives can be used to strengthen social interactions among workers leading to an improved work climate and more productive workers.

Our data, however, are not consistent with an interpretation of our findings based on *social incentive*s. First, we show that there is no interaction effect between altruism and firing treatments in explaining workers' production (see Table E.1 in Appendix E). This finding thus contrasts with the hypothesis derived from Rotemberg (1994). Also, our analysis of chat activities shows that in contrast to the prediction derived from Dur and Sol (2010) the magnitude

[16] Because in equilibrium no firing occurs, and because our data reports only a limited number of fired employees we cannot reliably estimate the claim that group production influence firing decisions more significantly in the *partial* and *minimal info* treatments compared to *complete info*.

of chatting activities does not differ between the *complete info* treatment (3.6% of available time is spent chatting) and the *social incentives* treatments (*minimal* and *partial info* where 3.1% of available time is spent chatting) (see Appendix F, p-value = 0.729 for the coefficient associated to the *social incentive* dummy that takes value 1 for the *minimal* and *partial info* treatments).[17] This result holds whether we consider only chatting activities between peers or whether we include chatting activities with the boss. Also, the proportion of time employees spent encouraging their peers did not differ across treatments (10.45% and 7.78% of messages in category 4 for *complete info vs partial plus minimal info*, respectively; 0.52% and 2.18% of messages in category 5 for *complete info vs partial plus minimal info*, respectively; and 0.91% and 0.52% of messages in category 7 for *complete info vs partial plus minimal info*, respectively).

It thus seems unlikely that either monitoring intensity or social incentives explain our findings. These interpretations of our results are also based on the idea that the *complete info* treatment may have underperformed because it demotivated workers because of excessive monitoring or a lack of *social incentives*. We study this possibility below by developing several measures of intrinsic motivation.

4.4.3. *Intrinsic motivation*

A final issue is whether our results could be explained in terms of differences in intrinsic motivation, of which we have two measures. First, we assess the production of workers in the last period of the experiment when firing threats were removed and when there were no more monetary incentives to produce. We find that last-period production is the highest in the *complete info* treatment in contrast with the idea that such treatment could negatively affect workers' morale. Last- period production was also higher in both the *partial* and *complete* treatments compared to the *baseline* whereas there was no difference in last-period production between the *minimal info* treatment and the *baseline*. Similar results are found if we consider working dedication instead of production levels (see Figs. G.1 and G.2 in Appendix G).

A second measure of intrinsic motivation was based on our end of experiment survey (see Appendix O.1). This measure was calculated as the ratio between the performance on the non-incentivized version of the 2-minute adding task (completed at the end of the experiment) and

[17] Chatting data correspond to the case of workers who have not been fired and can thus chat with other organizational members. We also do not include the last period data in our calculation as firing threats did not apply then. These analyses are included in Appendix F.

the performance on the incentivized version of the same task (completed before the experiment started). We do not find significant differences across firing treatments regarding our intrinsic motivation measure (Wilcoxon rank-sum test, p-values > 0.10 for all pairwise comparisons). Interestingly our intrinsic motivation measure is actually slightly (although not significantly) higher in our firing treatments compared to the *baseline* suggesting that firing threats, despite their strong incentive effects, do not generally have a negative effect on motivation.

5. Conclusions

We study the robustness of the incentive effect of firing threats to different monitoring technologies. Our aim was to assess whether firing threats could induce high levels of performance when information is scarce, which is a cornerstone assumption of the models demonstrating the optimality of firing schemes (e.g. Shapiro and Stiglitz, 1984).[18] By studying the performance of firing incentive schemes under different informational conditions, our work also responds to the call of Hart and Holmström (1986) to study the robustness of commonly-used incentive schemes.

We find that the incentive effect of firing threats was robust to cases in which bosses did not have access to any information about workers' individual performance. Also, the highest level of organizational production was obtained even when the only individual pieces information consisted in observing which activity employees were currently engaged in. Our findings can thus help account for the disconnection between pay and performance observed in compensation contracts (e.g. Baker, Jensen and Murphy, 1988). Our findings are also in line with the use of firing threats based on limited but uncontroversial and easily measured information about workers' dedication to their job such as absenteeism (Banerjee and Duflo, 2006; Duflo, Hanna and Ryan, 2012).

Our findings suggest that "a threat is a Threat" implying that workers may *perceive* apparently weak incentive schemes, based on limited information, to be strong. These potentially biased *perceptions* are not taken into account in current principal-agent models. By contrast, expectancy theory (Vroom, 1964) which is a major framework to study motivation in management disciplines stresses the essential role of the *perceived* link between pay and performance in assessing the effectiveness of incentive schemes. Because the literature on expectancy theory

[18] A weaker assumption would be to consider that even though precise individual information is observable by employers it is not verifiable (MacLeod and Malcomson 1989).

neither documents the nature of workers' misperceptions nor provides a formal model, taking these misperceptions into account in a principal-agent framework would be an interesting avenue for future research.

Appendix A. Additional tables and regression analyses

Table A.1

GLS regression with random effects for individual production (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	- .86	- 4.72	- 22.56	- 26.10	- 41.06	- 12.10
	(29.50)	(24.27)	(25.48)	(28.87)	(29.34)	(22.22)
Treatment+	- 6.88	30.74**	76.13***	36.37**	81.41***	44.70***
	(15.45)	(13.22)	(13.97)	(15.42)	(15.43)	(13.11)
Ability[19]	8.47***	5.45***	5.62***	7.43***	7.29***	4.49***
	(1.66)	(1.29)	(1.34)	(1.69)	(1.69)	(1.23)
Gender	19.61	27.03*	8.19	21.63	4.30	12.30
	(16.20)	(14.60)	(14.84)	(15.54)	(15.79)	(14.39)
Number of observations	419	423	426	422	425	429
R^2	0.1980	0.1591	0.2566	0.1756	0.2545	0.1424

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

[19] We compute ability as the number of correct answers in the mathematical task subjects do before the experiment.

Table A.2

GLS regression with random effects for working time (in seconds) (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	1108.56***	1123.65***	765.94***	1124.36***	764.85***	813.45***
	(38.19)	(39.01)	(97.51)	(39.22)	(99.34)	(98.67)
Treatment[+]	- 25.24	17.47	365.85***	41.03*	393.61***	349.62***
	(18.67)	(20.39)	(48.54)	(23.25)	(51.31)	(49.77)
Ability	- 2.44	- 4.21*	- 1.26	- 4.42*	- 1.07	- 2.62
	(1.92)	(2.22)	(3.85)	(2.67)	(4.23)	(4.41)
Gender	39.64*	20.30	- 1.44	21.36	- 2.65	- 20.79
	(23.88)	(22.61)	(49.80)	(29.04)	(52.99)	(51.66)
Number of observations	419	423	426	422	425	429
R^2	0.0416	0.0213	0.2591	0.0435	0.2998	0.2360

[+]Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table A.3

GLS regression with random effects for individual production for high ability workers (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	58.31	23.36	49.74	53.14	77.47	73.97*
	(46.17)	(37.02)	(43.84)	(45.84)	(49.07)	(42.57)
Treatment[+]	10.05	49.80**	95.93***	40.63	91.42***	48.52**
	(25.09)	(17.75)	(21.74)	(24.75)	(27.04)	(21.07)
Gender	71.43**	68.80***	26.33	51.22**	8.28	10.55
	(25.42)	(19.72)	(23.84)	(25.41)	(27.89)	(23.11)
Number of observations	219	227	222	226	221	229
R^2	0.0687	0.1596	0.2037	0.0687	0.1359	0.0611

[+]Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table A.4

GLS regression with random effects for individual production for low ability workers (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	128.29***	103.79***	51.82**	77.44***	26.15	17.28
	(30.62)	(26.21)	(25.02)	(25.29)	(23.97)	(25.68)
Treatment+	- .36	17.69	59.50***	15.65	56.92***	42.17***
	(17.16)	(18.60)	(17.11)	(17.99)	(16.32)	(16.05)
Gender	- 12.90	- 8.51	- 1.92	10.65	17.15	23.73
	(17.17)	(18.80)	(17.07)	(18.08)	(16.12)	(17.50)
Number of observations	200	196	204	196	204	200
R^2	0.0075	0.0142	0.1518	0.0161	0.16.78	0.0953

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table A.5

GLS regression with random effects for individual production for high ability workers (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	940.22***	1010.40***	715.16***	984.05***	702.84***	779.86***
	(105.94)	(79.51)	(120.76)	(111.19)	(137.38)	(122.39)
Treatment+	- 10.62	27.48	393.48***	38.98	405.91***	372.43***
	(30.31)	(32.13)	(70.64)	(39.51)	(77.42)	(73.30)
Gender	103.12**	39.91	- 1.41	54.77	6.61	- 43.54
	(48.51)	(40.17)	(73.40)	(59.90)	(85.09)	(74.45)
Number of observations	219	227	222	226	221	229
R^2	0.0843	0.1982	0.2899	0.0207	0.3156	0.2712

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table A.6

GLS regression with random effects for individual production for low ability workers (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	1176.25***	1122.69***	771.92***	1120.89***	771.72***	769.37***
	(22.85)	(41.19)	(120.68)	(21.47)	(119.25)	(126.98)
Treatment+	- 44.73***	8.94	333.36***	53.59**	377.90***	326.60***
	(15.73)	(25.30)	(66.70)	(21.47)	(65.27)	(69.51)
Gender	- 10.95	- 10.99	6.55	- 9.68	6.69	
	(15.73)	(24.96)	(63.71)	(20.82)	(62.21)	
Number of observations	200	196	204	196	204	200
R^2	0.0755	0.0024	0.2277	0.0445	0.2852	0.2023

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table A.7

Average (median) [standard deviation] individual production across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1-4	Period 5	Total
B subjects only	Firing Complete Info (including fired subjects)	3.30	3.86	3.97	3.74	3.72	2.72	3.51
		(3)	(3.25)	(4.25)	(3.5)	(3.5)	(2)	(3)
		[2.39]	[2.69]	[2.49]	[2.55]	[2.53]	[2.81]	[2.61]
	Excluding fired subjects	-	-	4.12	4.04	3.82	3.13	3.70
				(4.5)	(3.75)	(3.5)	(2)	(3.5)
				[2.41]	[2.41]	[2.49]	[2.79]	[2.55]
	Firing Partial Info (including fired subjects)	3.36	3.53	4.13	3.52	3.63	2.53	3.41
		(3)	(2.5)	(3.5)	(3.5)	(3)	(2)	(3)
		[3.01]	[3.10]	[2.98]	[2.71]	[2.95]	[2.61]	[2.92]
	Excluding fired subjects	-	-	4.29	3.88	3.76	2.84	3.59
				(3.5)	(4)	(3)	(2.5)	(3)
				[2.93]	[2.58]	[2.92]	[2.60]	[2.88]
	Firing Minimal info (including fired subjects)	2.37	2.93	3.58	2.87	2.94	1.18	2.59
		(1.75)	(2.5)	(3.25)	(3)	(2.5)	(0)	(2)
		[2.02]	[2.56]	[2.42]	[2.40]	[2.38]	[1.93]	[2.40]
	Excluding fired subjects	-	-	3.65	2.98	2.98	1.28	2.66
				(3.5)	(3)	(2.5)	(0)	(2)
				[2.39]	[2.38]	[2.37]	[1.98]	[2.39]
	Baseline	1.79	1.81	1.63	1.81	1.76	1.39	1.69
		(1)	(1)	(1)	(1)	(1)	(0.25)	(1)
		[2.21]	[2.58]	[1.92]	[2.56]	[2.32]	[2.12]	[2.28]
C subjects only	Firing Complete Info	4.17	4.92	5.17	4.08	4.58	4.25	4.52
		(4.5)	(3.5)	(4.75)	(3.5)	(4)	(4.25)	(4)
		[2.79]	[4.13]	[5.32]	[1.98]	[3.55]	[2.58]	[3.34]
	Firing Partial Info	2.58	3.33	2.67	4.33	3.23	2.42	3.07
		(2.5)	(2)	(2.75)	(3.25)	(2.5)	(.75)	(2.25)
		[2.40]	[2.54]	[1.33]	[3.14]	[2.39]	[3.18]	[2.52]
	Firing Minimal info	3.33	3.42	3.17	3.08	3.25	3.25	3.25
		(3.25)	(3.5)	(2.25)	(3.25)	(3.25)	(3.75)	(3.5)
		[1.33]	[1.46]	[2.21]	[1.46]	[1.55]	[2.16]	[1.64]
	Baseline	1.67	1.67	3.17	1.25	1.94	2.33	2.02
		(1.5)	(1.5)	(2.25)	(1.25)	(1.5)	(2.25)	(1.5)
		[1.86]	[1.72]	[3.33]	[1.04]	[2.14]	[1.78]	[2.05]

Table A.8

Average (median) [standard deviation] percentage of time subjects spent working across treatments.

	Treatment	Period 1	Period 2	Period 3	Period 4	Subtotal Periods 1-4	Period 5	Total
B subjects only	Firing Complete Info (including fired subjects)	92.58	93.95	87.66	85.59	89.94	65.79	85.11
		(96.62)	(97.19)	(95.73	(96.17)	(96.67)	(85.60	(96.05)
		[9.60]	[7.64]	[21.77]	[26.25]	[18.32]	[37.64]	[25.32]
	Excluding fired subjects	-	-	91.03	92.44	92.51	75.59	89.42
				(95.96)	(96.81	(96.84)	(89.26)	(96.41)
				[13.41]	[9.98]	[10.32]	[29.59]	[16.95]
	Firing Partial Info (including fired subjects)	94.22	94.41	92.88	87.66	92.29	68.58	87.55
		(96.83)	(97.99)	(98.21)	(97.92)	(97.68)	(89.82)	(97.49)
		[6.94]	[13.88]	[18.94]	[28.59]	[18.89]	[38.23]	[25.76]
	Excluding fired subjects	-	-	96.45	96.61	95.38	77.15	91.98
				(98.25)	(98.06)	(97.82)	(92.35)	(97.61)
				[4.63]	[4.52]	[8.52]	[31.19]	[16.98]
	Firing Minimal info (including fired subjects)	91.52	90.83	92.83	84.75	89.98	49.29	81.84
		(96.36)	(97.40)	(98.07)	(96.17)	(97.03)	(47.36)	(96.11)
		[13]	[15.03]	[15.17]	[23.60]	[17.35]	[37.79]	[28.07]
	Excluding fired subjects	-	-	94.58	88.01	91.25	53.23	84.02
				(98.08)	(96.63)	(97.18)	(49.86)	(96.35)
				[8.11]	[16.91]	[13.75]	[36.48]	[25.01]
	Baseline	79.78	63.33	55.01	50.41	62.13	42.26	58.16
		(88.13)	(72.72)	(61.93)	(54.41)	(72.42)	(32.27)	(64.32)
		[22.59]	[34.11]	[38.51]	[35.03]	[34.74]	[34.88]	[35.60]
C subjects only	Firing Complete Info	79.70	72.80	67.87	70.83	72.80	71.53	72.55
		(82.44)	(70.38)	(61.13)	(72.88)	(72.88)	(69.62)	(72.88)
		[15.53]	[18.82]	[21.24]	[24.69]	[19.49]	[20.54]	[19.35]
	Firing Partial Info	76.04	77.18	78.55	88.18	79.99	82.83	80.56
		()	()	()	()	()	()	()
		[15.06]	[16.22]	[16.25]	[6.04]	[14.00]	[10.74]	[13.29]
	Firing Minimal info	96.40	96.50	95.83	89.33	94.52	77.13	91.04
		(98.04)	(98.46)	(97.91)	(95.66)	(97.99)	(93.39)	(97.72)
		[4.28]	[3.34]	[5.47]	[13.09]	[7.72]	[38.31]	[18.72]
	Baseline	67.47	63.84	69.44	69.99	67.69	72.02	68.55
		(73.27)	(73.33)	(75.33)	(73.26)	(73.82)	(80.44)	(74.96)
		[25.59]	[34.34]	[26.40]	[24.32]	[26.17]	[26.38]	[25.81]

Table A.9

Period evolution of monitoring activities (% of total time).

Treatment	Proportion of total time (in %) *C* subjects spent monitoring	Period 1	Period 2	Period 3	Period 4
Firing Complete Info	14.58%	14.24%	15.68%	15.14%	13.25%
Firing Partial Info	10.40%	13.71%	9.93%	13.07%	4.88%
Baseline	12.93%	18.14%	12.47%	13.05%	8.06%

Table A.10

Tobit regression with random effects for monitoring time –in seconds- for periods 1 to 4.

	Firing Complete Info vs Partial Info	Firing Complete Info vs Baseline
Constant	117.94*** (46.02)	135.29*** (46.78)
Treatment[+]	53.91 (65.00)	35.93 (65.73)
Number of observations	*n* = 48 (3 left censored)	*n* = 48 (7 left censored)
Log likelihood (L)	L = -284.011 [Prob>χ2]=0.4069	L = -270.659 [Prob>χ2]=0.5846

[+]Treatment F is a dummy variable that takes value 1 for Treatment Firing Complete Info and 0 otherwise.

*p-Value<.10, **p-value <.05, and ***p-value <.01. (Standard deviation in parentheses)

Table A.11

Tobit regression with random effects for working time –in seconds- per period for workers.

	Firing Complete Info	Firing Partial Info	Firing Minimal info	Baseline
Constant	1110.97***	1130.70***	1098.22***	957.42***
	(25.40)	(25.24)	(32.60)	(56.36)
Period 2	-16.38	1.62	-8.29	-197.46***
	(32.21)	(33.82)	(40.28)	(46.28)
Period 3	-18.03	24.02	32.73	-308.06***
	(32.57)	(34.21)	(40.52)	(46.55)
Period 4	-2.04	25.81	-45.82	-366.95***
	(32.95)	(34.78)	(40.75)	(46.64)
Period 5	-205.76***	-207.57***	-461.72***	-469.43***
	(33.55)	(34.98)	(41.21)	(46.79)
Number of observations	*n* = 257 (0 right censored)	*n* = 257 (0 right censored)	*n* = 263 (0 right censored)	*n* = 270 (11 right censored)
Log likelihood (L)	L = -1701.259	L = -1700.363	L = -1803.411	L = -1858.559
	[Prob>χ2]<0.0001	[Prob>χ2]<0.0001	[Prob>χ2]<0.0001	[Prob>χ2]<0.0001

*p-Value<.10, **p-value <.05, and ***p-value <.01. (Standard deviation in parentheses)

Appendix B. Robustness analyses.

Table B.1

GLS regression with random effects for individual production (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	.63	1.34**	-.16	-.42	-.79	.10
	(.82)	(.64)	(.63)	(.74)	(.73)	(.57)
Treatment[+]	-.24	.84***	1.77***	1.04***	1.99***	1.02***
	(.38)	(.32)	(.36)	(.38)	(.37)	(.35)
Ability	.21***	.14***	.13***	.19***	.17***	.10***
	(.04)	(.03)	(.03)	(.04)	(.04)	(.03)
CRT	-.00	-.00	.19	-.00	.26	.22
	(.11)	(.10)	(.15)	(.12)	(.18)	(.16)
Gender	.33	-.55	.05	-.48	-.16	-.10
	(.40)	(.37)	(.38)	(.41)	(.42)	(.41)
Aheadness aversion	.25	-.28	.25	.07	.62	.22
	(.49)	(.47)	(.47)	(.49)	(.49)	(.43)
Behindness aversion	-.86**	-.46	-.73**	-.65	-.67*	-.46
	(.43)	(.42)	(.36)	(.43)	(.37)	(.35)
Number of observations	419	423	426	422	425	429
R^2	0.2170	0.1711	0.2893	0.1867	0.2925	0.1680

[+]Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.2

GLS regression with random effects for working time (in seconds) (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	1130.78***	1147.48***	824.70***	1132.59***	809.31***	824.45***
	(32.70)	(36.78)	(88.85)	(36.51)	(94.31)	(97.59)
Treatment+	-29.01*	22.86	374.59***	46.36**	398.54***	341.63***
	(17.12)	(21.52)	(47.51)	(22.89)	(50.55)	(51.19)
Ability	-2.75	-4.28*	-1.77	-5.23*	-1.28	-2.35
	(1.87)	(2.19)	(3.78)	(2.76)	(4.13)	(4.34)
CRT	2.62	-2.97	-10.15	4.70	-4.25	-12.75
	(6.16)	(7.91)	(16.57)	(8.57)	(19.40)	(19.82)
Gender	32.26	-12.04	9.43	16.63	-1.80	21.26
	(20.91)	(21.26)	(50.53)	(27.27)	(54.63)	(54.12)
Aheadness aversion	19.49	36.00	-40.17	59.07**	-28.19	-12.41
	(19.85)	(24.78)	(59.30)	(29.02)	(58.30)	(67.87)
Behindness aversion	-33.73*	-44.60	-75.60	-40.34	-66.89	-74.01
	(19.84)	(29.50)	(53.12)	(30.26)	(50.30)	(59.36)
Number of observations	419	423	426	422	425	429
R^2	0.0584	0.0361	0.2740	0.0632	0.3093	0.2453

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.3

GLS regression with random effects for individual production (all periods) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	.82	1.18*	-.27	-.07	-.63	.03
	(.76)	(.61)	(.61)	(.70)	(.68)	(.53)
Treatment[+]	-.17	1.02***	1.70***	1.14***	1.87***	.79**
	(.36)	(.30)	(.34)	(.35)	(.34)	(.32)
Ability	.19***	.12***	.13***	.16***	.16***	.10***
	(.04)	(.03)	(.03)	(.04)	(.03)	(.03)
CRT	-.01	-.01	.18	-.00	.27	.22
	(.12)	(.10)	(.15)	(.11)	(.03)	(.16)
Gender	.25	-.41	.05	-.29	-.23	-.03
	(.38)	(.34)	(.37)	(.38)	(.18)	(.39)
Aheadness aversion	.225	-.07	.39	.06	.54	.32
	(.46)	(.45)	(.46)	(.44)	(.47)	(.42)
Behindness aversion	-.78*	-.53	-.62*	-.62	-.56	-.45
	(.40)	(.38)	(.34)	(.38)	(.34)	(.33)
Number of observations	514	520	527	520	527	533
R^2	0.1718	0.1435	0.2679	0.1490	0.2667	0.1349

[+]Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.4

GLS regression with random effects for working time (in seconds) (all periods) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	1171.74***	1089.99***	748.69***	1124.60***	788.04***	745.33***
	(36.22)	(45.58)	(87.63)	(43.20)	(94.04)	(93.64)
Treatment+	-22.87	74.54***	390.49***	92.67***	407.71***	304.85***
	(19.21)	(25.60)	(48.47)	(23.71)	(50.07)	(51.39)
Ability	-4.83**	-5.09*	-1.03	-7.41**	-2.29	-1.91
	(2.12)	(2.76)	(3.83)	(2.90)	(4.10)	(4.37)
CRT	-3.82	-8.05	-13.44	7.39	.87	-7.20
	(8.45)	(9.75)	(18.05)	(9.53)	(20.64)	(20.70)
Gender	8.31	21.56	30.64	-12.25	-15.13	47.00
	(21.85)	(25.62)	(52.02)	(25.76)	(54.42)	(55.07)
Aheadness aversion	10.70	60.28*	-21.15	44.42	-49.26	2.13
	(24.78)	(30.97)	(64.94)	(31.53)	(62.22)	(72.47)
Behindness aversion	-25.16	-51.97	-58.81	-40.17	-59.93	-77.29
	(23.85)	(25.60)	(55.75)	(28.19)	(51.37)	(51.39)
Number of observations	514	520	527	520	527	533
R^2	0.0234	0.0418	0.2467	0.0578	0.2777	0.1619

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.5

GLS regression with random effects for individual production for high ability workers (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	1.49	3.59***	1.68**	1.69	2.30*	2.01***
	(1.66)	(.58)	(.65)	(1.49)	(1.30)	(.65)
Treatment[+]	.05	1.19***	1.98***	1.20**	2.09***	1.02*
	(.61)	(.42)	(.58)	(.61)	(.61)	(.60)
CRT	.33	.23	.47**	.10	.40	.28
	(.21)	(.14)	(.23)	(.20)	(.31)	(.24)
Gender	1.35*	-1.42**	-.08	1.01	-.45	-.01
	(.80)	(.56)	(.68)	(.68)	(.74)	(.63)
Aheadness aversion	.90	.28	.98	.62	1.74	.59
	(.83)	(.63)	(.80)	(.90)	(1.20)	(.91)
Behindness aversion	-.47	-.73	-.91	-.86	-.70	-.86
	(.86)	(.62)	(.64)	(.77)	(.73)	(.62)
Number of observations	219	227	222	226	221	229
R^2	0.1104	0.1982	0.2906	0.0893	0.2036	0.1078

[+]Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.6

GLS regression with random effects for individual production for low ability workers (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	3.58***	2.40***	1.29**	2.07***	.82	1.19*
	(.81)	(.54)	(.53)	(.66)	(.57)	(.65)
Treatment+	.04	.63	1.39***	.51	1.30***	1.08***
	(.44)	(.49)	(.47)	(.47)	(.38)	(.42)
CRT	-.04	-.07	.08	.04	.15	.16
	(.13)	(.13)	(.15)	(.14)	(.16)	(.19)
Gender	-.27	.24	.11	.20	.41	-.36
	(.42)	(.49)	(.44)	(.48)	(.43)	(.53)
Aheadness aversion	-.21	-.46	-.01	.05	.34	.29
	(.50)	(.50)	(.52)	(.39)	(.38)	(.42)
Behindness aversion	-.76*	-.31	-.46	-.58	-.88**	-.14
	(.44)	(.52)	(.41)	(.49)	(.38)	(.43)
Number of observations	200	196	204	196	204	200
R^2	0.0539	0.0297	0.1694	0.0347	0.2374	0.1118

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.7

GLS regression with random effects for working time for high ability workers (in seconds) (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	982.35***	1100.71***	918.74***	1025.60***	853.99***	928.46***
	(115.69)	(44.57)	(134.10)	(101.93)	(144.46)	(107.49)
Treatment[+]	-15.37	31.24	424.09***	56.59	449.23***	373.78***
	(28.78)	(30.58)	(63.89)	(38.84)	(73.81)	(68.24)
CRT	-4.79	-17.16	-23.08	1.75	-6.60	-15.30
	(11.42)	(10.80)	(20.36)	(14.34)	(27.41)	(23.04)
Gender	92.19**	-21.29	70.79	32.08	-54.72	89.35
	(45.71)	(34.08)	(84.80)	(50.74)	(88.51)	(73.73)
Aheadness aversion	10.99	42.89	33.01	70.58*	54.96	120.98
	(27.10)	(40.10)	(71.66)	(38.92)	(72.42)	(81.75)
Behindness aversion	-28.26	-91.17**	-192.33**	-76.79*	-161.06**	-203.91***
	(25.58)	(45.43)	(82.38)	(43.88)	(70.90)	(77.11)
Number of observations	219	227	222	226	221	229
R^2	0.0922	0.0749	0.3413	0.0633	0.3515	0.3250

[+]Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table B.8

GLS regression with random effects for working time for low ability workers (in seconds) (periods 1–4) across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	1163.57***	1098.4***	818.33***	1110.86***	796.63***	784.37***
	(20.54)	(47.82)	(81.03)	(38.62)	(113.12)	(137.18)
Treatment+	-52.79***	-10.82	320.74***	45.76**	374.39***	312.50***
	(16.17)	(26.18)	(68.97)	(20.06)	(63.46)	(70.18)
CRT	12.96**	14.67*	9.35	5.70	-.92	-11.47
	(5.34)	(7.67)	(25.39)	(6.61)	(29.70)	(36.87)
Gender	-18.03	-21.29	-19.75	-13.84	12.63	-33.29
	(15.64)	(34.08)	(62.99)	(19.90)	(74.68)	(85.75)
Aheadness aversion	2.90	42.89	-136.97	7.43	-130.86	-138.52
	(17.39)	(40.10)	(86.77)	(18.36)	(92.05)	(101.46)
Behindness aversion	2.20	-91.17**	3.46	29.75**	15.39	-22.78
	(15.90)	(45.43)	(75.41)	(15.09)	(82.46)	(85.81)
Number of observations	200	227	204	196	204	200
R^2	0.1240	0.0749	0.2560	0.0587	0.3126	0.2301

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Appendix C. Firing threats across ability level.

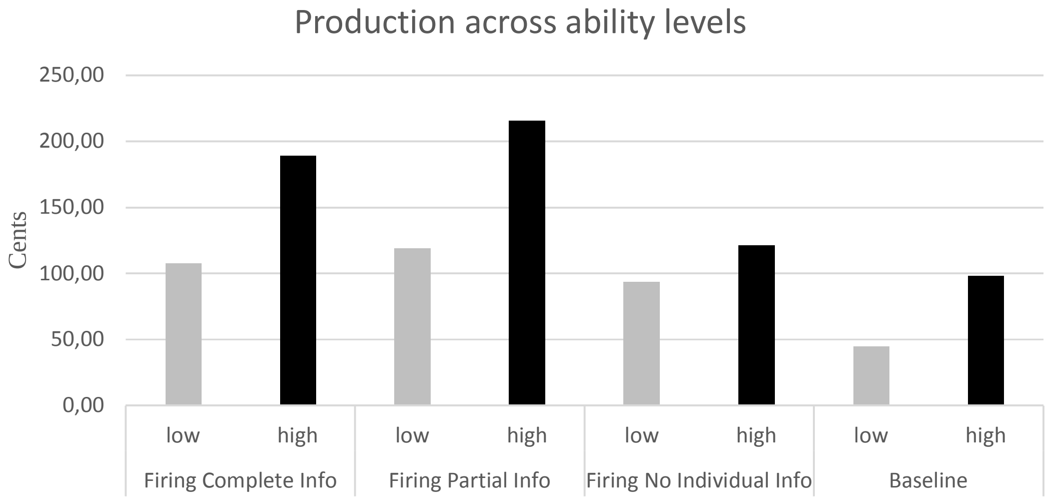

Fig. C.1 Employees' average total production across treatments by ability levels.

Appendix D. 5-minute analysis.

Table D.1

GLS regression with random effects for workers' production (all periods) across treatments where firing is allowed. Robust standard errors. Excluding fired workers.

	Firing Complete Info	Firing Partial Info
Constant	30.02***	24.50***
	(5.93)	(4.00)
Minute+	2.77	3.68**
	(2.12)	(1.53)
Watch++	.60	4.63
	(7.28)	(5.12)
Minute_Watch+++	- .60	- 1.09
	(2.83)	(2.23)
Number of observations	1028	1028
R^2	0.0032	0.0060

+Minute takes value 1 for the first 5 minutes of a period, 2 for next 5 minutes and so on until value 4.

++Watch is a dummy variable that takes the value 1 if a worker was observed in the previous 5-minute moment.

+++Minute_Watch id the interaction term between the previous variables.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Appendix E. Social incentives analysis.

Table E.1

GLS regression with random effects for workers' production (periods 1–4) across treatments where firing was allowed. Robust standard errors clustered by session. Excluding fired workers.

Constant	24.93
	(25.54)
SP[+]	- 7.32
	(20.08)
Altruism	6.22
	(6.70)
SP_Altruism	- 2.01
	(6.64)
Ability	7.31***
	(1.04)
Number of observations	632
R^2	0.1488

[+]SP is a dummy variable that takes value 1 for the minimal and partial info treatments and 0 for the complete info treatment. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Appendix F. Chat analysis.

Each chat message was assigned to one of thirty-three categories by two graduate students coding messages independently (see Table F.3). Then, we computed the Cohen's Kappa coefficient for each category to assess inter-rater agreement (see Table F.1).[20] We dropped category 18 and 19 from the analysis because they were empty and another five categories (categories 7, 20, 23, 24, and 33) because the corresponding Cohen Kappa test was not significant at a 5% significance level. These categories represented only 1.07% of the messages (see Figure F.1). The most represented category (33.82%) corresponds to distracting messages (e.g. jokes and stories). General and nonstrategic messages constituted the great majority (67.75%) of chat messages. We consider as general and nonstrategic messages the ones that were assigned to categories related to either presentation (category 1), distraction (categories 2 and 3) or general observations about the experiment (categories 27, 28, 29 and 30). Most of the strategic messages consisted in subjects stating their own performance (category 13, 6.24% of all messages) and encouraging others to produce (category 4, 4.67% of all messages).

We present disaggregate data at treatment level of the percentage of messages of each category (see Table F.2). We can observe that 44.31% of messages in the baseline treatment are related to category 2 (jokes and stories). This percentage is relatively high compared to Firing Complete Info (19.61), Firing Partial Info (15.83%), and Firing Minimal Info (20.87%). In relation to strategic messages the highest differences we find are related to categories 4 (Encouraging others to produce) and 13 (State your own performance). We observe that the percentage of messages in these categories is much higher in the Firing Complete Info treatment (10.65% for category 4, and 11.95% for category 13) compare to the baseline, Firing Partial Info and Firing Minimal Info treatments (2.27%, 4.38% and 3.41% respectively for category 4, and 5.32%, 1.88%, and 6.41% respectively for category 13).

In summary, chatting activities were mostly leisure activities. Indeed, similarly to Internet browsing, the average amount of time *B* subjects dedicated to chatting was significantly greater in the baseline treatment (31.54%) than in Firing Complete Info (4.71%), Firing Partial Info (3.89%), and Firing Minimal Info (10.08%).

[20] According to Landis and Koch (1977), Cohen Kappa coefficients between 0.4 and 0.6 correspond to a moderate agreement level and coefficients greater than 0.6 correspond to full agreement.

TABLE F.1 Inter-rater analysis of chat messages categorization.

Category	Agreement	Expected Agreement	Kappa	Standard Error	Z	Prob> Z
1	98.97%	92.97%	0.80	0.016	51.18	0
2	77.21%	54.82%	0.50	0.015	32.51	0
3	87.01%	82.69%	0.25	0.015	16.38	0
4	97.89%	91.1%	0.76	0.016	49.01	0
5	99.54%	98.63%	0.66	0.015	43.50	0
6	99.66%	98.60%	0.76	0.015	49.37	0
7	99.52%	99.52%	0.001	0.0093	-0.09	0.5371
8	97.94%	95.38%	0.55	0.015	35.74	0
9	99.39%	99.01%	0.39	0.014	27.39	0
10	99.66%	98.75%	0.73	0.015	47.73	0
11	99.15%	97.25%	0.69	0.015	45.53	0
12	99.81%	99.76%	0.20	0.016	12.78	0
13	96.39%	88.28%	0.69	0.015	44.84	0
14	99.85%	99.66%	0.57	0.014	40.59	0
15	99.10%	98.91%	0.17	0.014	12.35	0
16	99.73%	99.06%	0.72	0.016	46.04	0
17	99.76%	99.52%	0.50	0.014	34.98	0
18	100%	100%	0	0	0	0.5
19	100%	100%	0	0	0	0.5
20	100%	100%	0	0	0	0.5
21	97.60%	94.09%	0.59	0.015	38.86	0
22	97.87%	96.14%	0.45	0.015	30.38	0
23	99.61%	99.61%	0.002	0.015	-0.12	0.5493
24	99.3%	99.3%	0.002	0.017	-0.17	0.5680
25	99.76%	99.66%	0.28	0.015	19.08	0
26	98.59%	95.45%	0.69	0.015	45.12	0
27	99.66%	99.61%	0.12	0.010	12.08	0
28	97.72%	96.1%	0.42	0.015	27.40	0
29	96.05%	91.28%	0.55	0.015	36.98	0
30	84.22%	75.25%	0.36	0.014	26.14	0
31	95.73%	91.68%	0.49	0.014	34.58	0
32	98.79%	98.74%	0.04	0.010	3.98	0
33	99.52%	99.52%	0.001	0.009	-0.09	0.5373

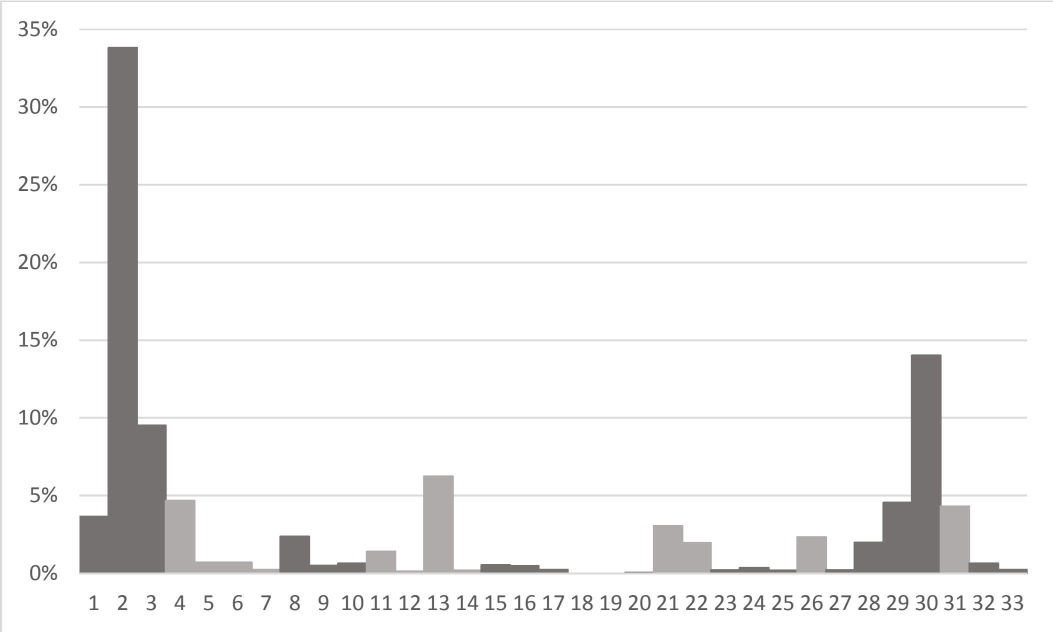

Figure F.1 Histogram of categorization of messages for all treatments.

Table F.2 Percentage of categories by treatment.

Category	Baseline	FC	FP	FM
1	3.53	2.34	4.38	3.83
2	44.31	19.61	15.83	20.87
3	8.47	3.64	9.79	14.26
4	2.27	10.65	4.38	3.41
5	0.61	0.52	1.46	0.73
6	0.21	4.42	0.84	0.00
7	0.07	0.91	0.21	0.31
8	1.50	3.51	6.46	5.89
9	0.42	1.04	0.83	0.93
10	0.15	2.60	0.84	0.83
11	0.92	1.04	0.21	3.72
12	0.04	0.13	0.00	0.11
13	5.32	11.95	1.88	6.41
14	0.17	0.00	0.00	0.62
15	0.57	0.13	0.83	0.73
16	0.00	2.73	2.09	0.73
17	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00
19	0.00	0.00	0.00	0.00
20	0.00	0.00	0.00	0.11
21	3.97	2.86	1.88	1.35
22	1.02	1.69	2.92	3.41
23	0.02	0.00	1.46	0.83
24	0.09	0.13	0.00	0.11
25	0.00	0.00	0.00	1.45
26	2.38	3.51	6.46	1.76
27	0.09	0.00	0.00	0.00
28	1.96	3.51	2.92	1.76
29	4.28	3.51	9.79	4.44
30	13.51	15.59	14.79	14.26
31	3.30	3.64	7.29	6.31
32	0.79	0.39	1.04	0.52
33	0.11	0.00	1.46	0.42
Total messages	2398	385	240	484

Table F.3 Categories for chat messages.

Group Category	Category Number	Category
Social interaction	1	Greetings (Hello/Goodbye)
	2	Distracting others (jokes, stories)
	3	Personal chat (talking about likes and dislikes)
Positive feedback and help	4	Encouraging others to produce
	5	Thanking other for their cooperative behavior
	6	*C* give positive feedback about *B* contributions
	7	Help others complete the task
Discouragements	8	Discouraging others to produce
	9	Asking others what is the point of producing anything
	10	*C* give negative feedback about *B* contributions
Performance evaluation and comparison	11	Ask others' performance on the task
	12	*B* asks *C* about his/her own relative performance on the task
	13	State your own performance
	14	B talks to C about other B subjects' performance
Pay /firing threats	15	*B* threatening *C* not to produce anything
	16	*C* threatening others to fire them if they do not produce enough
	17	*C* telling *B* they will be paid based on their relative production
	18	*C* telling *B* they will be paid based on how much time they spent working instead of being online
	19	*C* telling all *Bs* they will all be paid the same if they achieve a certain level of total production
	20	*C* telling all *Bs* they will all be paid the same regardless of performance
Complaints about firing/supervision strategy/pay	21	Complaints about the supervision of the *C* subject
	22	Complaints about the firing/pay strategy of the *C* subject
Comments on firing/supervision/pay strategy	23	Suggesting/stating Firing strategy
	24	Suggesting/stating Supervising strategy
	25	Comments on effectiveness of firing policy
Envy	26	*B* envying the *C* subject
Non-strategic comments on the experiment	27	Ask others for help and hints to complete the task
	28	General comments about the experiment and its goals
	29	Specific comments on how earnings are calculated
	30	Other specific comments on the experiment
Influence and manipulation	31	Influencing *C* subject
Fairness	32	Negative comments on fairness of firing / pay policy
	33	Positive comments on fairness of firing / pay policy

Table F.4

GLS regression with random effects for chat time (in seconds) (periods 1–4) across treatments where firing was allowed. Robust standard errors clustered by session. Excluding fired workers.

Constant	3.59***
	(1.27)
SP+	- .57
	(1.64)
Number of observations	632
R^2	0.0026

+SP is a dummy variable that takes value 1 for the minimal and partial info treatments and 0 for the complete info treatment. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Appendix G. Intrinsic motivation

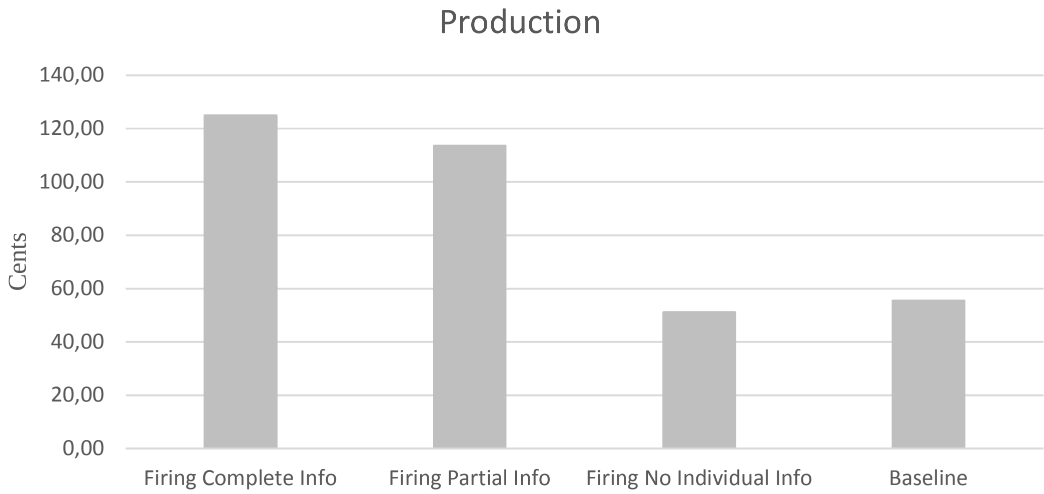

Fig G.1 Workers' average production in period 5 across treatments. Excluding fired workers.

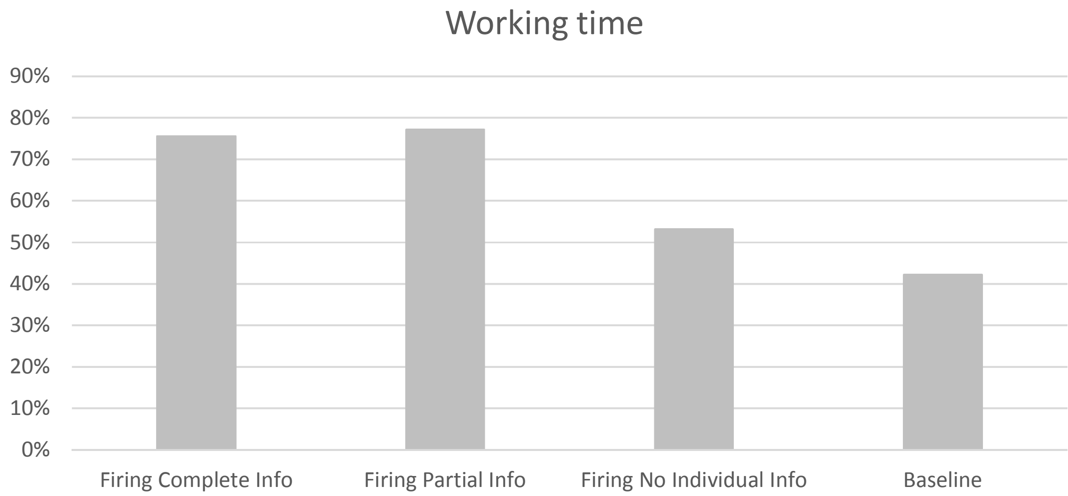

Fig G.2 Workers' average working time (%) in period 5 across treatments. Excluding fired workers.

Table G.1

OLS regression for workers' production in period 5 across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	57.65**	19.16	- 13.26	30.08	- 7.91	17.49
	(26.53)	(28.61)	(24.65)	(22.25)	(20.21)	(20.22)
Treatment+	4.87	72.64***	65.18***	63.99***	61.15***	-5. 24
	(21.56)	(19.40)	(19.15)	(19.10)	(18.43)	(16.06)
Ability	4.07**	2.17	4.76**	1.43	4.39***	2.63*
	(1.80)	(1.77)	(1.73)	(1.28)	(1.44)	(1.44)
Number of observations	95	97	101	98	102	104
R^2	0.0465	0.1443	0.1847	0.1111	0.1495	0.0324

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

Table G.2

OLS regression for workers' working time (in seconds) in period 5 across treatments. Robust standard errors. Excluding fired workers.

	Firing Complete Info vs. Partial Info	Firing Complete Info vs. Minimal info	Firing Complete Info vs. Baseline	Firing Partial Info vs. Minimal info	Firing Partial Info vs. Baseline	Firing Minimal info vs. Baseline
Constant	97.54***	70.97***	44.13***	80.70***	50.61***	47.05***
	(7.62)	(10.15)	(9.12)	(9.32)	(9.01)	(9.74)
Treatment+	0.80	23.06***	33.45***	22.04***	34.50***	11.09
	(6.05)	(6.62)	(6.38)	(6.60)	(6.58)	(7.02)
Ability	- 1.48**	- 1.20*	- .13	- 1.86***	- .58	- .33
	(.52)	(.63)	(.53)	(.59)	(.53)	(.58)
Number of observations	95	97	101	98	102	104
R^2	0.0734	0.1385	0.2109	0.1840	0.2265	0.0262

+Treatment is a dummy variable that takes value 1 for the first treatment in the comparison and 0 otherwise. Excluding fired subjects.

*p-Value<0.1, **p-value<.0.05, and ***p-value<.0.01. (Standard deviation in parentheses)

References

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011): "Reference points and effort provision," *American Economic Review*, 101(2), 470–492.

Baker, G., M. Jensen and K. Murphy (1988) "Compensation and Incentives: Practice vs. Theory," *Journal of Finance* 43(3), 593-616.

Bandiera, O., I. Barankay, and I. Rasul (2010), "Social Incentives in the Workplace," Review of Economic Studies 77(2), 417-458.

Banerjee A. and E. Duflo (2006): "Addressing absence," *Journal of Economic Perspectives*, 20, 117-132.

Bartling, B., E. Fehr, M. A. Maréchal, and D. Schunk (2009): "Egalitarianism and Competitiveness," *American Economic Review: Papers & Proceedings*, 99(2), 93–98.

Bartling, B. and F. von Siemens (2010), "The Intensity of Incentives in Firms and Markets: Moral Hazard With Envious Agents," *Labour Economics* 17 (3), 598-607.

Becker, G. S., and G. J. Stigler (1974): "Law Enforcement, Malfeasance, and Compensation of Enforcers," *The Journal of Legal Studies*, 3(1), 1-18.

Bénabou, R., and J. Tirole (2003): "Intrinsic and Extrinsic Motivation," *Review of Economics Studies*, 70, 489-520.

Besley, T and M. Ghatak (2005): "Competition and Incentives with Motivated Agents," *American Economic Review*, 616-636.

Bolton, P. and M. Dewatripont (2005): "Contract Theory," *MIT Press*, Cambridge, MA.

Cameron, A. C and D. Miller (2011): "Robust Inference with Clustered Data," *Handbook of Empirical Economics and Finance*, ed. by A. Ullah and D. E. A. Giles. Boca Raton, CRC Press, 1-28.

Charness, N. and J. Campbell (1988): "Acquiring skill at mental calculation in adulthood: a task decomposition," *Journal of Experimental Psychology: General* 117, 115-129.

Charness, G., R. Cobo-Reyes, N. Jimenez, J. A. Lacomba, and F. Lagos (2017): "Job security and long-term investment: An experimental analysis," *European Economic Review*, 95, 195-214.

Charness, G., D. Masclet, and M.C. Villeval (2014): "The Dark Side of Competition for Status," *Management Science*, 60(1), 38-55.

Chiappori P. and B. Salanié (2000): "Testing Contract Theory: A Survey of Some Recent Work," in: M. Dewatripont, L.P. Hansen, S. Turnovski, Advances in Economic Theory, Eighth World Congress of the Econometric Society, Cambridge University Press.

Corgnet, B., J. Gómez-Miñambres, and R. Hernán-González (2105): "Goal Setting and Monetary Incentives: When Large Stakes Are Not Enough," *Management Science*, 61(12), 2926-2944.

Corgnet, B., Gómez-Miñambres, J., and R. Hernan-Gonzalez (2016): "Goal Setting in the Principal-Agent Model: Weak Incentives for Strong Performance," Working paper GATE 2016-28.

Corgnet, B., R. Hernán-González, and S. Rassenti (2015a): "Firing Threats: Incentive Effects and Impression Management," *Games and Economic Behavior*, 91, 97-113.

Corgnet, B., R. Hernán-González, and S. Rassenti (2015b): "Peer pressure and moral hazard in teams: Experimental evidence," *Review of Behavioral Economics*, 2(4), 375-403.

Corgnet, B., R. Hernán-González, and E. Schniter, (2015): "Why real leisure really matters: incentive effects on real effort in the laboratory," *Experimental Economics*, 18(2), 284-301.

Deci E. L., R. Koestner, and R. M. Ryan (1999): "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation," *Psychological Bulletin*, 125, 627–668.

Deci, E. L. (1971): "Effects of Externally Mediated Rewards on Intrinsic Motivation," *Journal of Personality and Social Psychology*, 18(1), 105-115.

Deci, E. L. (1975): Intrinsic Motivation. New-York: Plenum Press.

Deci, E. L., and R. M Ryan (1985): Intrinsic Motivation and Self-Determination in Human Behavior. (1 ed.). Springer.

Deci, E. L., and R. M. Ryan (2014): "The Importance of Universal Psychological Needs for Understanding Motivation in the Workplace," In M. Gagne (Ed.). Oxford University Press.

Deci, E. L., R. M. Ryan, M. M. Gagne, D. R. Leone, J. Usunov, and B. P. Kornazheva (2001): "Need satisfaction, motivation, and well-being in the work organizations of a former eastern bloc country: A cross-cultural study of self-determination," *Personality and Social Psychology Bulletin*, 27(8), 930-942.

Dickinson, D.(1999): "An experimental examination of labor supply and work intensities," *Journal of Labor Economics*, 17, 638–670.

Dickinson, D., and M. C. Villeval (2008) “Does Monitoring Decrease Work Effort?: The Complementarity Between Agency and Crowding-Out Theories,” *Games and Economic Behavior*, 63, 56-76.

Dohmen, T., and A. Falk (2010): “You get what you pay for: incentives and selection in the education system,” *Economic Journal*, 120 (546), 256–271.

Dohmen, T., and A. Falk (2011): “Performance Pay and Multi-Dimensional Sorting: Productivity, Preferences and Gender,” *American Economic Review*, 101 (2), 556-590.

Duflo E., R. Hanna, and S. Ryan (2012): “Incentives work: getting teachers to come to school,” American Economic Review 102, 1241–1278.

Dur, R. and J. Sol (2010): “Social Interaction, Co-Worker Altruism, and Incentives,” *Games and Economic Behavior* 69(2), 293-301.

Eriksson, T., A. Poulsen, and M.C. Villeval (2009): “Feedback and incentives: experimental evidence,” *Labour Economics*, 16 (6), 679–688.

Falk, A., and A. Ichino (2006): “Clean evidence on peer effects,” *Journal of Labor Economics*. 24 (1), 39–58.

Falk, A., and D. Huffman (2007): “Studying labor market institutions in the lab: minimum wages, employment protection, and workfare,” *Journal of Institutional and Theoretical Economics*, 163 (1), 30–45.

Falk, A., D. Huffman, and W.B. Macleod (2015): “Institutions and contract enforcement,” *Journal of Labor Economics*, 33(3), 571-590.

Gneezy, U., and J.A. List (2006): “Putting behavioral economics to work: testing for gift exchange in labor markets using field experiments,” *Econometrica*, 74 (5), 1365–1384.

Goerg, S. and S. Kube (2012): “Goals (th)at Work- Goals, Monetary Incentives, and Workers’ Performance,” Working Paper Series of the Max Planck Institute for Research on Collective Goods Bonn, 2012/19.

Holmström, B. (1979): “Moral Hazard and Observability,” *Bell Journal of Economics*, 10(1), 74-91.

Holmström, B. and P. Milgrom (1991) “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design,” *Journal of Law, Economics and Organization* VII, 24-52.

Kandel, E., and E. Lazear (1992): "Peer Pressure and Partnerships," *Journal of Political Economy* 100, 801-817.

Klein, B., and K. B. Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, University of Chicago Press, 89(4), 615-41.

Kosfeld, M. and S. Neckermann (2011): "Getting More Work for Nothing? Symbolic Awards and Worker Performance," *American Economic Journal: Microeconomics* 3(3), 86-99.

Laffont, J. J. and D. Martimort (2002): The theory of incentives: The principal-agent model. Princeton University Press: Princeton, N.J.; Oxford.

Levin, J. (2003) "Relational Incentive Contracts," *American Economic Review* 93(3), 835-57.

Lumesse (2013): "The Impact of Complexity," http://www.lumesse.com/

MacLeod, W. B., and J. Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57(2), 447-80.

Malachowski, D., 2005, Wasted time at work costing companies billions. San Francisco Chronicle, July 11.

http://www.sfgate.com/cgi-bin/article.cgi?f=/g/a/2005/07/11/wastingtime.TMP.

Manso, G. (2011): "Motivating Innovation," *Journal of Finance* 66, 1823-1869.

Nalbantian, H. and A. Schotter (1997): "Productivity Under Group Incentives: An Experimental Study," *American Economic Review*, 87(3), 314-341.

Rotemberg, J. (1994). "Humans relations in the workplace," *Journal of Political Economy*, 102: 684-717.

Shapiro, C., and J. Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74(3), 433-44.

Van den Steen, E. (2010) "Interpersonal Authority in a Theory of the Firm," *American Economic Review* 100(1), 466-490.

Vroom, V. (1964): "Work and Motivation," San Francisco, CA: Jossey-Bass.

Wu, G., C. Heath, and R. Larrick (2008): "A Prospect Theory Model of Goal Behavior," mimeo University of Chicago.

Online Appendix.

Appendix O1. Additional tables

We detail below the tests which were completed by subjects as part of the one-hour survey conducted at the lab in which the experiment was performed. This survey was conducted at the beginning of the year, about six months before completion of Study 1.

Summation skills

The instructions for this task were as follows. Instructions:

This task consists in adding five one-digit numbers. During a period of 2 minutes you can solve as many problems as you want to. An example of the sum problem is displayed below. Next to the display, there is an input box and an O.K. button. You will have to enter the result into the box (only integer numbers are allowed) and then click on the O.K. button. For each sum problem that you solve correctly, you will receive 10 cents. If you enter a wrong result and click O.K., a message 'Last answer was not correct.' will be displayed. You will be informed about the number of problems you have solved correctly (on the right hand side of the screen). The time remaining in seconds will be displayed in the upper left corner of the screen.

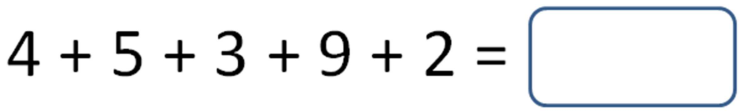

Figure O1.1: Example of Adding Task question.

Intrinsic motivation

To measure intrinsic motivation, we assess the extent to which people performed on the previous adding task in the absence of any monetary incentives. We then computed the intrinsic motivation score as the ratio between one’s performance on the task without incentives and one’s performance on the task in the presence of monetary incentives. The incentive version of the task was presented first, and the non-incentivized version of the task was presented at the end of the survey.

Social motives

Participants made six choices between two possible allocations of money between themselves and another anonymous participant with whom they were randomly matched. In each experimental session (typically composed of 12subjects), two participants and one of the six decisions were selected at random for payment. The choice of the first participant in the selected decision was used to allocate payoffs between the two participants. All decisions were anonymous. The first four decisions used the exact same payoffs as in Bartling et al. (2009). Decisions 5 and 6 were added by Corgnet et al. (2015).

All the allocation decisions are described in Table C.3. Option A always yielded an even distribution of money ($10 to both the self and the other participant) whereas option B yielded uneven payoffs. For each decision, we show in parentheses the envy/compassion parameter associated to choosing the egalitarian and non-egalitarian options (i.e. options A and B) and in square brackets the proportion of subjects who chose each option. Note that the model parameters associated to Decisions 1-4 are the same as in Study 1, except for the fact that in Decision 4 the threshold for the envy parameter is now 0.125 instead of 0.5.

Table O1.1. Decisions in the social preferences task (Study 1). For each option, we display the payoff for the decision-maker and the recipient, the associated model parameters (in parentheses) and the % of subjects choosing it (in square brackets).

Decision #	Option A self, other		Option B self, other	
1	$10,$10	[80%]	$10,$6	[20%]
2	$10,$10	[33%]	$16,$4	[67%]
3	$10,$10	[49%]	$10,$18	[51%]
4	$10,$10	[34%]	$11,$19	[66%]
5	$10,$10	[48%]	$12,$4	[52%]
6	$10,$10	[89%]	$8,$16	[11%]

The *altruism* index is calculated as the number of times one chooses Option A for decisions 1, 2 and 5 and Option B for decisions 3, 4 and 6. The higher the index the more likely a person values the other person's payoff positively.

Extended cognitive reflection test (CRT):

Taken from Frederick (2005):

(1) A bat and a ball cost $1.10 in total. The bat costs a dollar more than the ball. How much does the ball cost? ____ cents
[Correct answer: 5 cents; intuitive answer: 10 cents]

(2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? ____ minutes
[Correct answer: 5 minutes; intuitive answer: 100 minutes]

(3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? ____ days
[Correct answer: 47 days; intuitive answer: 24 days]

Taken from Toplack et al. (2014):

(4) If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? _____ days
[correct answer: 4 days; intuitive answer: 9]

(5) Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? ______ students
[correct answer: 29 students; intuitive answer: 30]

(6) A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How much has he made? _____ dollars
[correct answer: $20; intuitive answer: $10]

(7) Simon decided to invest $8,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon has: a. broken even in the stock market, b. is ahead of where he began, c. has lost money
[correct answer: c; intuitive response: b]

Appendix O2. Model

Firing incentive schemes

We focus on pure-strategy Bayesian equilibria that induce the non-intrinsically motivated agents to exert effort. These are the equilibria that maximize the principal's welfare as long as we assume that $w < 1$. We also assume that $w > \frac{\alpha_i^{NI}+\beta_i}{2}$ so that it is not only optimal for the principal but also efficient for non-intrinsically motivated agents to exert effort.

i) *Minimal information* (y). In the case in which the principal only observes total production of agents, he or she can only incentivize workers by threatening to fire a proportion *f* of workers at the end of a given period whenever total production falls below a certain threshold ($\bar{y}$).

In that context, there exists a pure-strategy efficient equilibrium in which no agents are fired and non-intrinsically motivated agents exert effort until period T_f. After that period, no agents are fired and only intrinsically-motivated workers exert effort. Equilibrium strategies are as follows.

- Principals fire a proportion f of workers at the end of each of the first T_f periods if total production falls below $\bar{y}$. No agents are fired otherwise. After T_f periods nobody is fired.[21] Non-intrinsically motivated agents will exert effort for the first T_f periods.
- Assuming non-intrinsically motivated agents believe the firing policy of the principal applies, they will exert effort in the first T_f periods as long as: $f(T - T_f)w > \frac{\alpha_i^{NI}+\beta_i}{2}$.
- Principals will follow their firing policy as long as it is worthwhile inducing effort from non-intrinsically motivated agents which is the case as long as: $w < 1$. Also, the principal does not have an incentive to unconditionally fire agents whenever $w \leq \pi_I$.[22]

ii) *Partial information* (y, l_i). In the case of partial information, the conditions to obtain the efficient equilibrium described in i) are more easily achieved. This is so because the principal's strategy can be enhanced by threatening to fire people if $l_i > 0$. In that case, agents deviating will face a sure dismissal unless they decide they stay at their work station (without working, $e_i^{NP} = 1$) instead of undertaking the leisure activity. The efficient equilibrium described in i) thus now holds when the following looser condition is satisfied: $f(T - T_f)w > \frac{\alpha_i^{NI}}{2}$.

iii) *Complete information* (y, l_i, q_i). In that case, the efficient equilibrium can be sustained by firing those employees who produce nothing, regardless of total production. In that case, the efficient equilibrium always holds because by assumption $(T - T_f)w > \frac{\alpha_i^{NI}+\beta_i}{2}$.

[21] A similar equilibrium holds when considering that a proportion of agents is always fired after period T_f.

[22] The case in which $w > \pi_I$ would require an additional condition to ensure that principals do not fire workers regardless of total production.